

Wells GR, Weichmann F, Sloan KE, Colvin D, Watkins NJ, Schneider C.
[The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain.](#)

Nucleic Acids Research 2017

DOI: <http://dx.doi.org/10.1093/nar/gkw1344>

Copyright:

© The Author(s) 2017. Published by Oxford University Press on behalf of Nucleic Acids Research. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

DOI link to article:

<http://dx.doi.org/10.1093/nar/gkw1344>

Date deposited:

13/01/2017



This work is licensed under a [Creative Commons Attribution 4.0 International License](http://creativecommons.org/licenses/by/4.0/)

The ribosome biogenesis factor yUtp23/hUTP23 coordinates key interactions in the yeast and human pre-40S particle and hUTP23 contains an essential PIN domain

Graeme R. Wells, Franziska Weichmann, Katherine E. Sloan, David Colvin, Nicholas J. Watkins and Claudia Schneider*

Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK

Received October 12, 2016; Revised December 21, 2016; Editorial Decision December 22, 2016; Accepted December 22, 2016

ABSTRACT

Two proteins with PIN endonuclease domains, yUtp24(Fcf1)/hUTP24 and yUtp23/hUTP23 are essential for early pre-ribosomal (r)RNA cleavages at sites A0, A1/1 and A2/2a in yeast and humans. The yUtp24/hUTP24 PIN endonuclease is proposed to cleave at sites A1/1 and A2/2a, but the enzyme cleaving at site A0 is not known. Yeast yUtp23 contains a degenerate, non-essential PIN domain and functions together with the snR30 snoRNA, while human hUTP23 is associated with U17, the human snR30 counterpart. Using *in vivo* RNA–protein crosslinking and gel shift experiments, we reveal that yUtp23/hUTP23 makes direct contacts with expansion sequence 6 (ES6) in the 18S rRNA sequence and that yUtp23 interacts with the 3' half of the snR30 snoRNA. Protein–protein interaction studies further demonstrated that yeast yUtp23 and human hUTP23 directly interact with the H/ACA snoRNP protein yNhp2/hNHP2, the RNA helicase yRok1/hROK1(DDX52), the ribosome biogenesis factor yRrp7/hRRP7 and yUtp24/hUTP24. yUtp23/hUTP23 could therefore be central to the coordinated integration and release of ES6 binding factors and likely plays a pivotal role in remodeling this pre-rRNA region in both yeast and humans. Finally, studies using RNAi-rescue systems in human cells revealed that intact PIN domain and Zinc finger motifs in human hUTP23 are essential for 18S rRNA maturation.

INTRODUCTION

Eukaryotic ribosomal (r)RNAs are processed from an initial 35S (*Saccharomyces cerevisiae*) or 47S (*Homo sapiens*) precursor (pre-rRNA) by a series of endonucleolytic cleavages followed by exonucleolytic trimming, which results in the concomitant removal of external (5' ETS, 3' ETS) and internal (ITS1, ITS2) transcribed spacer sequences (Figure 1 and Supplementary Figure S1) (1). The early pre-rRNA cleavages, at sites A0, A1 and A2 in yeast and A', A0, A1/1 and 2a/E in humans, are critical for 18S rRNA maturation and require the small subunit (SSU) processome, a large ribonucleoprotein complex (2). The U3 small nucleolar (sno)RNA, a key component of the SSU processome, base-pairs with the 5' ETS and 18S rRNA sequences to guide the formation of the conserved central pseudoknot, which is an essential feature of the 40S ribosomal subunit (2,3).

Many of the factors involved in ribosome biogenesis are essential for pre-rRNA cleavages, which has made the identification of the nuclease activities responsible for these cleavage events difficult. The PIN (Pit N-terminus) endonucleases yUtp24/Fcf1 (hUTP24 in humans) and yNob1 (hNOB1 in humans) are critical for 18S rRNA processing. Early pre-rRNA cleavages at three sites, A0, A1/1 and A2/2a, require the presence of yUtp24/hUTP24 and the PIN endonuclease domain of the protein is essential for cleavages at A1/1 and A2/2a (4–7). The yeast and human proteins also specifically cleave site A2 *in vitro* and yeast yUtp24 was shown to crosslink to both the U3 snoRNA and close to the A1 cleavage site in the pre-rRNA *in vivo* (7). yNob1/hNOB1 functions later in the 18S rRNA processing pathway and catalyzes the removal of the final part of ITS1 from the 3' end of the 18S rRNA (site D/3) in the cytoplasm (5,8–11). Only the nucleases that cleave at the A0

*To whom correspondence should be addressed. Tel: +44 191 208 7708; Fax: +44 191 208 7424; Email: Claudia.Schneider@ncl.ac.uk

Present Address:

Franziska Weichmann. Lehrstuhl für Biochemie I, Fakultät für Biologie und Vorklinische Medizin, Universität Regensburg, Regensburg, 93053, Germany.
Katherine E. Sloan. Universitätsmedizin Göttingen, Institut für Molekularbiologie, Göttingen, 37073, Germany.

and the metazoan-specific A' sites have yet to be assigned in the 18S rRNA maturation pathway.

In addition to yUtp24/hUTP24, one other PIN domain protein, yUtp23 (hUTP23 in humans), is essential for early cleavages at A0, A1/1 and A2/2a (4,12,13). However, yeast yUtp23 contains only two of the four conserved amino acids in the catalytic site and the two conserved PIN domain residues are not essential for yUtp23 function (4,14). In addition, the protein contains a conserved CCHC Zinc finger. yUtp23 binds nucleotides 745–859 in the 18S rRNA *in vitro* and specifically associates with the snR30 H/ACA snoRNP *in vivo* (14,15). The snR30 snoRNA is essential for the A0, A1 and A2 cleavages and SSU processome assembly (16,17), and is needed for the integration of yUtp23 into the pre-ribosome (15). Conversely, snR30 does not require yUtp23 for pre-ribosome association, but yUtp23 is essential for snR30 release from the complex (15). The internal loop in the 3' hairpin of snR30 base-pairs with two elements (rm1 and rm2) in the 18S rRNA expansion sequence 6 (ES6) and is expected to play an important role at this region during pre-SSU maturation (18,19). Interestingly, yUtp23 may be needed to establish base-pairing of snR30 with the 35S pre-rRNA (15). A number of other factors interact with snR30 and/or ES6 including yRok1 (snR30 and ES6 (20)), yRrp7 (ES6 (21)), yUtp24 (ES6 (7)) and yRrp5 (snR30 and ES6 (22)). Release of snR30 from the pre-ribosome requires the RNA helicase yRok1, which directly interacts with yRrp5 (22–24).

Considerably less is known about UTP23 in *Metazoa* but similar to the yeast protein, it is required for pre-rRNA cleavages in the 5' ETS and ITS1 (at sites A0, 1 and 2a in humans and mice) (12,13) and is also associated with the human homologue of snR30, U17 (15). However, hUTP23 has three of the four key acidic amino acids of the PIN domain, together with the Zinc finger motif, suggesting that it may be an active nuclease in human cells, but why such activity would be required in humans, but not in yeast, remains unclear.

Here, we present a combination of *in vivo* and *in vitro* approaches to determine the role of yUtp23/hUTP23 in both yeast and human ribosome biogenesis. *In vivo* RNA-protein crosslinking studies (CRAC) generated a transcriptome-wide RNA binding profile for yeast yUtp23, which provides new insights into its relationship with snR30 and the ES6 region of the 18S rRNA. We also performed *in vitro* assays using recombinant proteins to determine RNA and protein interaction partners of yeast yUtp23 and human hUTP23 in the SSU processome. Finally, we established RNAi-rescue systems in HEK293 cells to study the effect of mutant hUTP23 on pre-rRNA cleavage in the human system.

MATERIALS AND METHODS

Yeast strains and methods

The *S. cerevisiae* strain expressing genomically encoded, C-terminal HTP-tagged (His₆-TEV-protA) yUtp23 under the control of its endogenous promoter (Supplementary Table S1) was constructed by standard methods. Cultures were grown at 30°C in medium containing 2% glucose and 0.67% nitrogen base.

CRAC and data analysis

Actively growing yeast cultures in SD medium (OD₆₀₀ ~0.5) were UV-irradiated in a 1.2 m metal tube for 100 s at 254 nm to generate RNA–protein crosslinks. The CRAC method (Supplementary Figure S2A) was performed as described in (25,26). Illumina sequencing data were aligned to the yeast genome using Novoalign (<http://www.novocraft.com>). Sequencing data from this publication were analyzed as previously reported (27) and submitted to the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>, identifier GSE87238).

Cloning and mutagenesis

The open reading frames of the protein genes listed in Supplementary Tables S2 and S3 were amplified from yeast genomic DNA or human cDNA adding restriction sites and cloned into pET100 vectors (Invitrogen). The constructs were either used directly or sub-cloned into different protein expression vectors to purify affinity-tagged recombinant proteins from *Escherichia coli* using standard techniques or used for *in vitro* translation with [³⁵S] methionine (TNT, Promega).

A human hUTP23 cDNA construct (generated by IDT) containing a C-terminal His₈-PP (PreScission protease recognition site)-2xHA (hemagglutinin) tag was amplified by PCR and cloned into the pcDNA5/FRT/TO vector (Invitrogen) for protein expression under the control of a tetracycline-inducible promoter. The coding sequence of hUTP23 was altered to render it resistant to the siRNAs used to deplete the endogenous mRNA (Supplementary Table S4). Point mutations were generated by site-directed mutagenesis using overlapping primers (Supplementary Table S3) and confirmed by sequencing.

In vitro RNA binding electromobility shift assay (EMSA)

Recombinant proteins were expressed and purified as described previously (28). RNA substrates were transcribed in the presence of [³²P]-αUTP, using plasmid constructs or PCR products with T7 promoter sequences as template (Supplementary Tables S2 and S3). RNA binding assays were performed using trace amounts of radiolabeled RNA in 10 mM Tris/HCl pH 7.6, 75 mM NaCl, 2 mM dithiothreitol, 100 ng/μl bovine serum albumin, 0.8 units/μl RNasin, 4.5% glycerol, 0.05% Tween20 and 500 ng/μl *E. coli* tRNA. A total of 10 μl or 15 μl reactions containing 100–5000 nM protein were incubated for 10 min at 30°C, followed by 10 min incubation on ice, and 2 μl or 3 μl native agarose loading dye (30% glycerol and 0.3% Orange G (w/v)) was added. Products were resolved on 4% polyacrylamide/1x TBE (5% glycerol) native gels and visualized using a Typhoon FLA9000 PhosphorImager.

Protein–protein interaction studies

GST-bait proteins were immobilized on glutathione sepharose pre-blocked with bovine serum albumin and incubated with His-tagged recombinant proteins or [³⁵S] *in vitro* translates for 1 h at 4°C in Buffer NB (20 mM Tris/HCl pH 7.6, 150 mM NaCl, 8.7% glycerol and 0.1% Tween20). The beads were washed five times with buffer

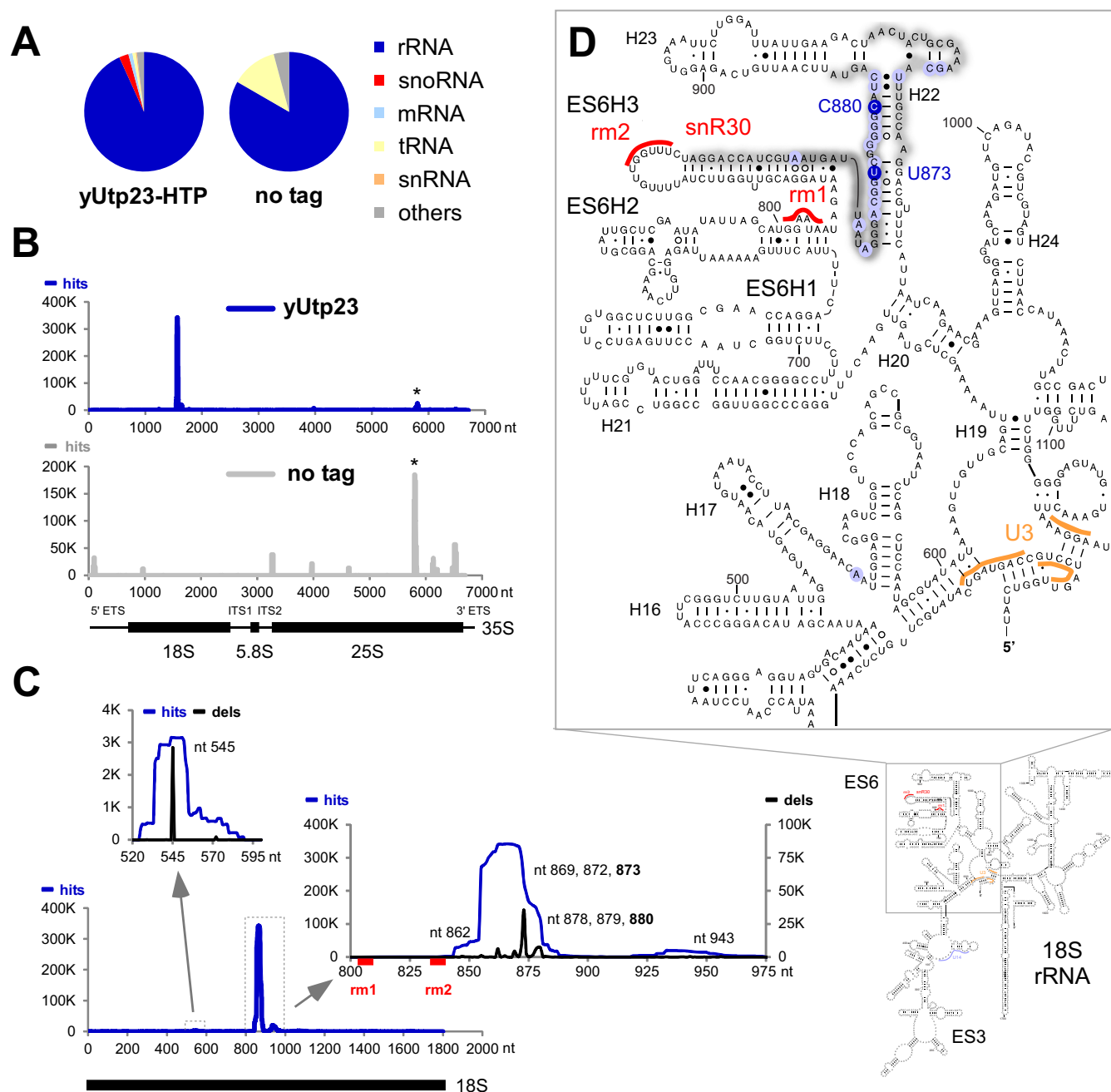


Figure 1. RNA crosslinking sites of yeast yUtp23 on the (pre-) ribosomal RNA. Crosslinked and subsequently trimmed RNAs were purified from yeast strains expressing yUtp23-HTP or a non-tagged control strain and used to generate cDNA libraries. Illumina sequencing data were mapped to the yeast genome using Novoalign. Normalized data for one representative yUtp23-HTP data set and the no tag control are plotted as reads per million mapped sequences ('hits'). (A) Transcriptome-wide binding profiles. A total of 2 277 820 mapped reads were recovered for the yUtp23-HTP data set and 7446 reads for the no tag control. Pie charts illustrate the proportion of all reads mapped to functional RNA classes (indicated on the right). (B) yUtp23 crosslinking profile on the primary ribosomal RNA transcript. Sequences were aligned with the rDNA (RDN37-1) encoding the 35S pre-ribosomal RNA. The frequency of recovery in the yUtp23 data set (blue) or the no tag control (gray) is plotted as total reads (hits) for each individual nucleotide. The positions of the mature 18S, 5.8S and 25S rRNAs are indicated by thick lines. A common CRAC contaminant at the 3' end of the 25S rRNA is marked by an asterisk (*). ETS: external transcribed spacer; ITS: internal transcribed spacer. (C) yUtp23 crosslinks on the 18S rRNA. Hits (blue): total reads; deletions (black; dels): mutations and microdeletions representing precise binding sites. Prominent microdeletion peaks around nt 545 (left insert) and in the 18S ES6 region (right insert) are labeled. The position of the mature 18S rRNA and the snR30 binding sites in the 18S ES6 region (rm1 and rm2) are indicated by a thick black line or red boxes, respectively. (D) Predicted secondary structure of the mature 18S rRNA in *S. cerevisiae*. yUtp23 crosslinking sites are marked on the sequence and gray shades indicate peak height. Microdeletion peaks (see panel C) are highlighted by shaded blue circles. Binding sites for the snoRNAs snR30 in the expansion sequence 6 (ES6) region (rm1 and rm2, red) and U3 around the central pseudoknot (orange) are also indicated.

NB to remove unbound material. Retained proteins were separated by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie staining, immunoblotting or by using a PhosphorImager, respectively.

Cell culture and RNAi

hUTP23-His₈-PP-2xHA-pcDNA5 constructs or the empty pcDNA5 plasmid were transfected into FLP-In T-Rex HEK293 cells as described by the manufacturer (Invitrogen) and stably transfected cells were cultured according to standard protocols. Expression of exogenous proteins was induced by addition of tetracycline (WT: 0.1–1 ng/μl; D31N: 50–100 ng/μl; C103A: 100–200 ng/μl). For RNAi-mediated depletion of endogenous proteins, cells were transfected with siRNA duplexes (Supplementary Table S4) using Lipofectamine RNAiMAX transfection reagent (Invitrogen) and harvested after 72 h.

RNA analysis

TRI reagent (Sigma-Aldrich) was used to extract RNA from HEK293 cell pellets. For northern blot analysis, 2 μg of total RNA was separated on 1.2% glyoxal-agarose gels, transferred to nylon membrane and hybridized with [³²P] 5' end radiolabeled oligonucleotides (Supplementary Table S3). Results were visualized with a Typhoon FLA9000 PhosphorImager and quantified using the ImageQuant software (GE Healthcare).

Immunofluorescence

HEK293 cells expressing HA-tagged hUTP23 proteins were grown on coverslips and induced with tetracycline for 72 h before being fixed with phosphate buffered saline (PBS) containing 4% paraformaldehyde. Immunofluorescence analysis was performed as described in (5). Briefly, cells were permeabilized using 0.2% Triton (v/v) and then incubated with primary and secondary antibodies (Supplementary Table S5) diluted in 10% fetal calf serum, PBS, 0.1% Triton (v/v) and washed with PBS, 0.1% Triton (v/v). In the final step, cells were washed with PBS containing DAPI (4',6'-diamidino-2-phenylindole, 1:10,000) and mounted onto a slide using Mowiol. Images were obtained using a Zeiss Axiovert 200 microscope with Plan-Apochromat x100 1.4NA objective, Axiovision software and an AxioCam monochrome camera and processed in Photoshop (Adobe).

Immunoprecipitation

Immunoprecipitation experiments were performed with sonicated whole-cell extracts as previously described (29), with the exception that anti-HA antibody-coupled agarose beads were used. Co-precipitated RNA was extracted, separated by denaturing polyacrylamide electrophoresis and analyzed by northern blotting using [³²P] 5' end radiolabeled oligonucleotides (Supplementary Table S3).

RESULTS

Yeast yUtp23 crosslinks to the 18S ES6 region and to the snR30 snoRNA

In yeast, yUtp23 is essential for 18S rRNA maturation (4) and required for dissociation of the snR30 small nucleolar RNP from pre-ribosomal particles (15), but its exact role within the SSU processome is unclear. We therefore applied *in vivo* RNA-protein crosslinking (26) to determine yUtp23 RNA binding sites (see Supplementary Figure S2A for overview of the CRAC procedure). We first constructed a yeast strain expressing genomically encoded, C-terminal HTP-tagged (His₆-TEV-protA) yUtp23 protein under the control of its endogenous promoter. The affinity-tagged yUtp23 protein supported wild-type growth and actively growing yeast cultures were UV-irradiated as described (25). Purification of yUtp23 proteins was verified by western analysis (Supplementary Figure S2B) and co-purified crosslinked RNA fragments were isolated and analyzed as outlined in Supplementary Figure S2A. We reproducibly detected yUtp23 RNA crosslinking sites in four independent CRAC experiments. Results from the two largest data sets (>2 million mapped reads each) are presented in Figures 1 and 2 and Supplementary Figure S4.

Figure 1A shows the transcriptome-wide RNA binding profile of yUtp23 compared to a non-tagged control strain. The majority of reads (>93%) in all four yUtp23 data sets were mapped to rRNA sequences, with lower numbers of hits in snoRNAs (on average ~2%). Importantly, no snoRNAs hits were detected in the non-tagged control data set. Within the primary 35S pre-rRNA transcript, yUtp23 crosslinks were predominately found in the 18S rRNA sequence (Figure 1B). This crosslinking profile is in agreement with the known function of yUtp23 in 18S rRNA maturation and was not seen with the non-tagged control strain.

The main crosslinking peaks in the 18S rRNA sequence (Figure 1C) are located between 18S nt 844–887 and 18S nt 929–953 of the eukaryotic ES6 or directly downstream of this region (helix 22 and 23, respectively) with some reads also found between 18S nt 534–554 (helix 17 and helix 18). To validate the main crosslinking site in the ES6 region, we performed *in vitro* RNA binding studies with recombinant GST-tagged yUtp23 protein expressed in *E. coli* and a radiolabeled 18S rRNA fragment encompassing the crosslinking sites (18S nt 775–963) (Figure 1D and Supplementary Figures S2C and S2D). Consistent with previously published data (14), the yUtp23 protein (but not the GST-only control) bound to this RNA fragment, albeit with very low efficiency. However, the formation of the RNP complex was specifically decreased by addition of a cold competitor RNA of the same sequence, but not an unrelated RNA fragment of similar size (18S nt 1022–1146). This result indicates specific binding of yUtp23 to this region of the rRNA.

During preparation of the cDNA libraries used for sequencing, microdeletions and/or mutations are often introduced at the site of crosslinking and can therefore be exploited to determine precise protein contact sites (30). Major sites of microdeletion within the main 18S rRNA peaks are highlighted in the inserts shown in Figure 1C. Positioning of the predominant yUtp23 reads and microdele-

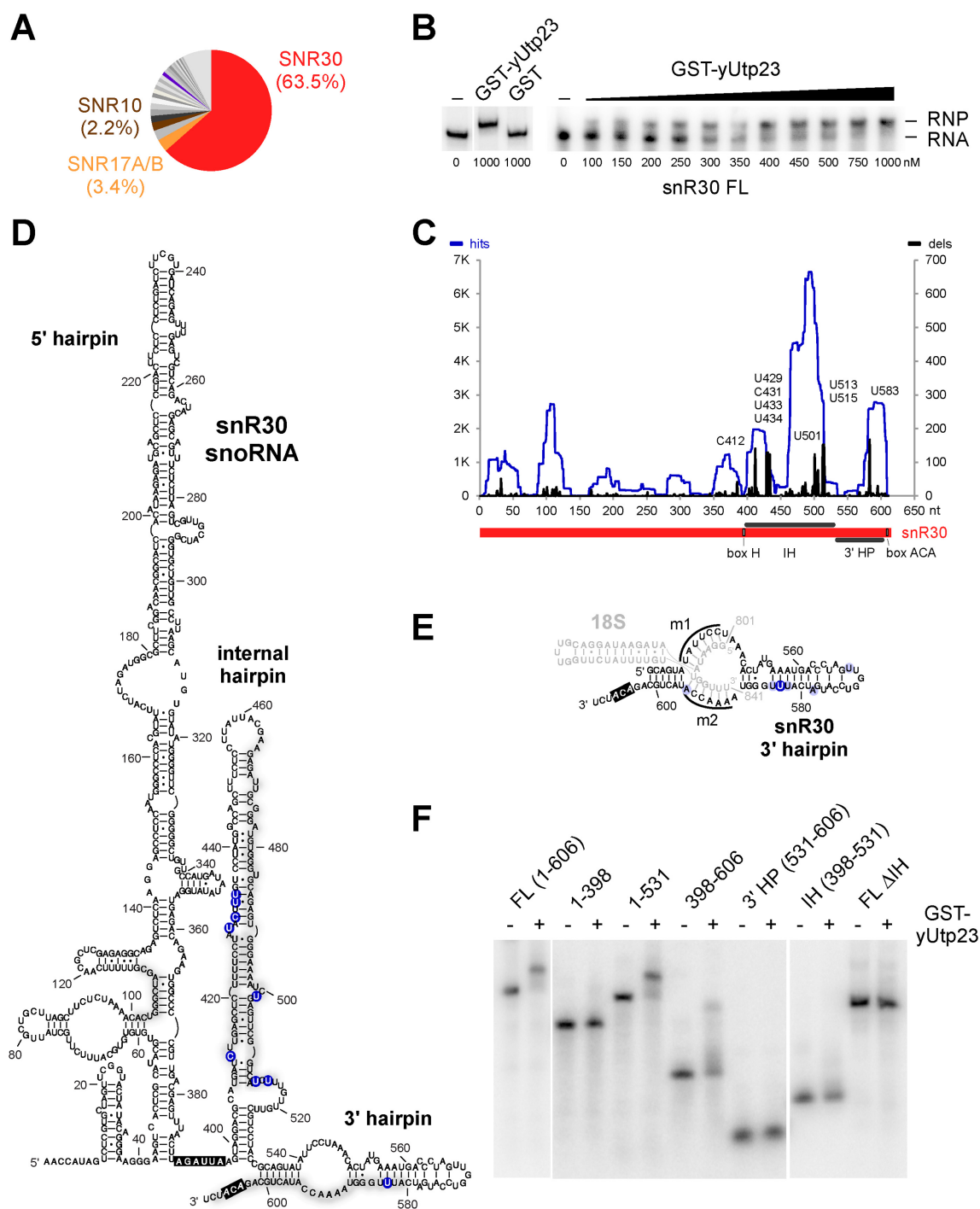


Figure 2. Yeast yUtp23 contacts the internal and 3' terminal hairpins of the snR30 snoRNA. (A) Pie chart of snoRNA reads from yUtp23 crosslinking data sets. The top 20 hits from two combined yUtp23 data sets (see Supplementary Figure S4B for more detail) are represented as percentages of all reads mapped to snoRNAs. Significant and reproducibly enriched RNAs with essential, non-modifying roles in 18S maturation (snR30, U3 and snR10) are highlighted. (B) Electromobility shift assay (EMSA) showing the binding of GST-yUtp23 or GST (1000 nM, left) or GST-yUtp23 (100–1000 nM, right) to *in vitro* transcribed radiolabeled snR30 RNA. Free RNA and RNP complexes were separated on 4% native polyacrylamide gels and visualized using a PhosphorImager. (C) yUtp23 crosslinking profile on the snR30 snoRNA. Hits (blue): total reads; deletions (black); dels: mutations and microdeletions representing precise binding sites. Prominent microdeletion peaks in the internal hairpin (IH) and 3' terminal hairpin (3' HP) are labeled. The positions of the H and ACA boxes within the snR30 RNA are indicated. (D) Predicted secondary structure of the snR30 snoRNA in *S. cerevisiae* (adapted from (18)). yUtp23 crosslinking sites are marked on the sequence and grey shades indicate peak height. Predominant microdeletion peaks (see panel C) are highlighted by shaded blue circles. The H and ACA boxes are shown in black. (E) A model of the interaction between the 18S ES6 binding sites (m1 and m2) in the 3' terminal hairpin (3' HP) of snR30 (black) and the 18S ES6 region (nt 801–841, grey). yUtp23 reads (grey) and microdeletion sites in the 3' HP (light and dark blue circles) are indicated. Shades of blue represent peak height (see panel C). The ACA box is highlighted in black. (F) EMSA showing the binding of GST-yUtp23 (1300 nM) to full-length (FL) snR30 or snR30 RNA fragments. Free RNA and RNP complexes were analyzed as in panel B. IH: internal hairpin; FL ΔIH: full-length snR30 lacking the internal hairpin.

tion sites on the predicted structure of the mature 18S rRNA (Figure 1D) revealed that the main yUtp23 binding sites are in close proximity, but not overlapping with the known snR30 binding sites rm1 (18S nt 801–806) and rm2 (18S nt 836–842) in the ES6 region. Direct yUtp23 protein–RNA contacts within this region were also detected by primer extension analysis on non-digested RNA that was crosslinked to and co-purified with yUtp23 (Supplementary Figure S3A). Notably, similar protein–RNA contacts were also observed between hUTP23 and the human ES6 region when HEK293 cells, expressing an affinity-tagged version of hUTP23, were used for crosslinking and primer extension analysis (Supplementary Figure S3B). This indicates that the interaction between yUtp23/hUTP23 and the 18S ES6 region is evolutionarily conserved.

We next examined the yUtp23 crosslinking sites on snoRNAs (Figure 2 and Supplementary Figure S4). For this, the data for the 20 snoRNAs, to which the maximum number of reads mapped in two independent yUtp23 data sets, were combined and are represented as percentages of reads mapped to all yeast snoRNAs (Figure 2A). Remarkably, 63.5% of snoRNA hits mapped to a single box H/ACA snoRNA, snR30. Significant proportions of reads were also reproducibly mapped to two of the three other snoRNAs involved in 18S rRNA processing (Supplementary Figure S4A), the box C/D snoRNA U3/snR17A/B (3.4%) and the H/ACA snoRNA snR10 (2.2%), whereas reads corresponding to the third, the box C/D snoRNA U14/snR128, were only recovered in one of the analyzed data sets (Supplementary Figure S4B). Hits were also seen for the box C/D snoRNAs snR190 (2.7%) and other modification snoRNAs (0.4–1.8%). The substantial enrichment of snR30 over other snoRNAs suggests that the previously observed relationship between yUtp23 and snR30 involves a direct protein–RNA interaction. Consistent with this, recombinant GST-tagged yUtp23 protein specifically bound to *in vitro* transcribed snR30 RNA, with an estimated dissociation constant (K_D) value of ~ 300 nM (Figure 2B). yUtp23 crosslinking hits and microdeletions could also be mapped to the U3/snR17A/B, snR10 and U14/snR128 snoRNA sequences suggesting direct protein–RNA contacts *in vivo* (Supplementary Figure S4C). However, stable interactions with these RNAs were not detected *in vitro* (data not shown) and might therefore require the presence of the pre-rRNA or other factors within the pre-ribosomal particle.

The positions of yUtp23 crosslinking reads and microdeletion sites on the snR30 sequence (Figure 2, panels C and D) suggest that yUtp23 mainly contacts the 3' portion of the snoRNA. Most reads and deletions mapped to the internal hairpin (IH, nt 398–531) and the 3' terminal hairpin (3' HP, nt 532–601) of snR30, whereas fewer reads were found in the 5' hairpin. The high number of yUtp23 crosslinking sites mapping to the internal hairpin was surprising given that this snR30 region was shown to be dispensable for growth *in vivo* (18). Interestingly, the majority of the precise crosslinking sites on the snR30 3' HP did not overlap with the snR30 sequences (m1 and m2) that can engage in base-pairing interactions with the 18S rRNA (Figure 2E). Instead, most microdeletions were located within the distal region of the 3' HP, which was previously pro-

posed to be a putative snR30-specific snoRNP protein binding site (19).

To investigate the importance of the snR30 IH and 3' HP with respect to yUtp23 association, we designed a series of snR30 fragments and tested them for binding to recombinant yUtp23 proteins *in vitro* (Figure 2F). yUtp23 efficiently bound the full-length (FL) snR30 (nt 1–606). In agreement with the CRAC data suggesting significant contacts to the internal hairpin (IH), binding of yUtp23 was severely reduced in the absence of the IH (FL Δ IH). yUtp23 exhibited no binding to the snR30 5' hairpin alone (nt 1–398), while binding was restored when the IH (nt 1–531) was included. Moreover, neither the IH (nt 398–531), nor the 3' HP (nt 531–606) alone were sufficient for binding, whereas combination of both (nt 398–606) enabled binding, but at a reduced level. The data suggest that the internal hairpin and 3' hairpin in snR30 are both needed for yUtp23 binding *in vitro* and that elements in all three hairpins are necessary for maximally efficient binding.

We conclude that yUtp23 primarily binds to the 3' portion of the snR30 snoRNA and adjacent to the snR30 base-pairing sites in the mature 18S rRNA sequence. These findings support a direct role for yUtp23 in snR30 function and release.

Yeast yUtp23 and human hUTP23 directly bind proteins that interact with the 18S rRNA ES6 region and snR30/U17

In recent years, several *in vivo* RNA-protein crosslinking studies in yeast have highlighted the 18S ES6 region as a binding platform for early acting SSU synthesis factors, which contact the 18S rRNA in close proximity to the described yUtp23 crosslinking sites (Figure 3A). In the case of the PIN domain endonuclease yUtp24 (7) and the RNA helicase yDhr1 (31), binding to the ES6 region only represented a secondary pre-rRNA contact, while their main crosslinking sites were located within the pseudoknot region. Consistent with this, snoRNA crosslinks for both proteins were mainly mapped to U3, which forms extensive base-pairing interactions with the central pseudoknot. Two other factors, the RNA helicase yRok1 (20) and the RNA-binding protein yRrp7 (21), were almost exclusively crosslinked to the ES6 region. While both proteins are functionally linked to snR30, yRok1 was shown to crosslink to the SSU-associated snoRNAs snR30, U3 and U14, and, to a lesser extent, to snR10, whereas yRrp7 exhibited a strong association with snR10. Lastly, the CTD domain of the pre-ribosomal 'compaction factor' yRrp5 also crosslinked to the ES6 region and analysis of full-length yRrp5 revealed hits to all SSU-associated snoRNAs (U3, snR30, snR10 and U14) (22). Many of these SSU synthesis factors are conserved in the human system and are therefore predicted to engage in similar interactions at the human ES6 region or with the U17 snoRNA, the human snR30 counterpart, but their exact contact sites or interaction partners are not known.

To better understand the role of yUtp23/hUTP23 within this 'binding platform' for ribosome biogenesis factors in yeast and humans, we established an *in vitro* system using recombinant proteins (Supplementary Figures S2 and S5) and [35 S] *in vitro* translated proteins to analyze protein–protein

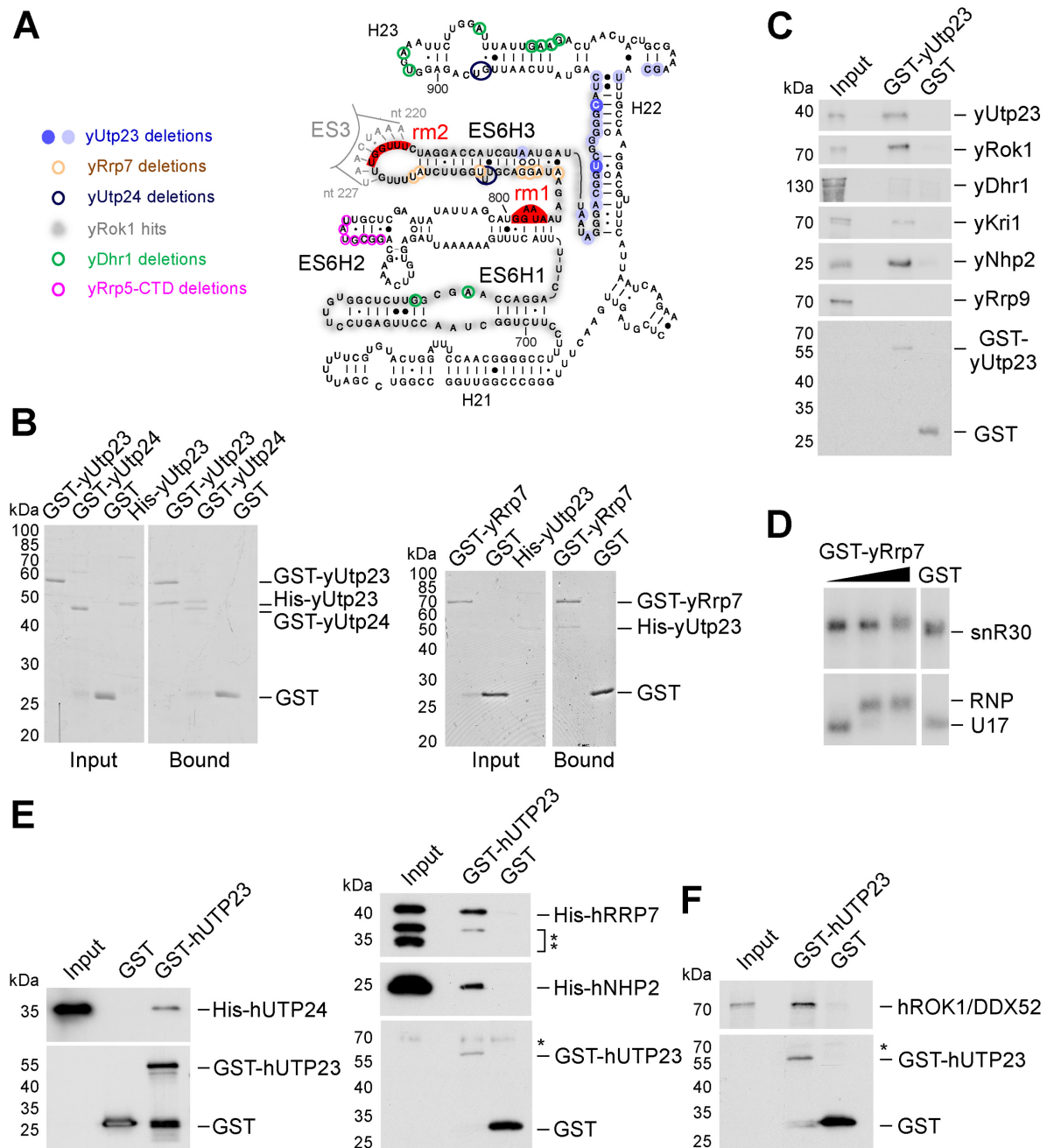


Figure 3. Yeast yUtp23 and human hUTP23 interact with early pre-40S factors contacting the 18S rRNA ES6 region. (A) Predicted secondary structure of the mature yeast 18S ES6 region with binding sites of crosslinked ribosome biogenesis factors. Precise crosslinking sites (microdeletions) of yUtp23 (shaded blue circles indicating peak height, see Figure 1C and D) and other published factors (yUtp24 (7), dark blue; yRrp7 (21), brown; yDhr1 (31), green and the yRrp5-CTD (22), pink) are highlighted by circles. yRok1 crosslinking regions (total reads, (20) are shaded in grey. Binding sites for snR30 (rm1 and rm2) are indicated in red. (B) Recombinant GST-tagged yeast yUtp23, yUtp24, yRrp7 or free GST were immobilized on glutathione sepharose and incubated with protein-A-His-tagged yUtp23. Bound material was eluted under denaturing conditions, separated by SDS-PAGE and visualized by Coomassie staining. A total of 10% of the input material was loaded. (C) Immobilized GST-tagged yeast yUtp23 or free GST was incubated with proteins generated by *in vitro* translation in the presence of [³⁵S] methionine. Bound material was treated as in panel B and analyzed by a PhosphorImager. 2% of the input material was loaded. (D) EMSA showing the binding of GST-yRrp7 (170, 1700 and 3300 nM) or GST (3300 nM) to *in vitro* transcribed radiolabeled snR30 (top) or human U17 (bottom) RNA. RNA and RNP complexes were analyzed as in Figure 2B. (E) Recombinant GST-tagged human hUTP23 or free GST was immobilized on glutathione sepharose and incubated with N-terminally His-tagged hUTP24, hRRP7 or hNHP2. Bound material was eluted under denaturing conditions, separated by SDS-PAGE and transferred to nitrocellulose membrane. Proteins were analyzed by immunoblotting using antibodies specific for the His-tag (upper panels) or the GST-tag (lower panel), respectively. Double asterisks: C-terminally truncated forms of the His-tagged hRRP7 protein. The single asterisk denotes a non-specific protein recognized by the anti-GST antibody (lower panel). Left panel (hUTP24): 10% of the input material was loaded. Right panel (hRRP7, hNHP2): 5% of the input material was loaded. (F) Immobilized GST-tagged human hUTP23 or free GST was incubated with hRok1/DDX52 protein generated by *in vitro* translation in the presence of [³⁵S] methionine. Bound material was treated as in panel C and analyzed by a PhosphorImager. One percent of the input material was loaded. Asterisk: non-specific protein recognized by the anti-GST antibody.

interactions between ES6-interacting factors (Figure 3, panels B, C, E and F). In these experiments, yeast or human recombinant GST-tagged proteins or free GST were immobilized on glutathione sepharose, incubated with His-tagged recombinant or *in vitro* translated proteins and washed repeatedly to remove non-specifically bound factors. Retained material was eluted under denaturing conditions, separated by SDS-PAGE and visualized by Coomassie staining (panel B), PhosphorImaging (panels C and F) or immunoblotting (using anti-His or anti GST tag antibodies, panel E), respectively.

We firstly found that yeast yUtp23 interacted with itself (panels B and C), suggesting that, similar to other PIN domain proteins (32), it might act as a multimer. Interestingly, we also detected strong interactions between yUtp23 and the PIN domain endonuclease yUtp24 (panel B), which has recently been reported to be directly responsible for the A1 and A2 site cleavages (6,7). This raises the exciting possibility that yUtp23 might act as a recruitment factor or, alternatively, a chaperone for yUtp24, to modulate yUtp24 catalytic activity until it is correctly positioned within the SSU processome. We did not, however, observe interactions between yUtp23 and yDhr1 (panel C) or yRrp5 (not shown).

yUtp23 also specifically interacted with yRrp7 (panel B) and yRok1 (panel C). In yeast, snR30 is required for the stable association of yRrp7 to pre-ribosomes but not vice versa (21). We therefore speculate that yRrp7 contacts the pre-rRNA either during snR30 base-pairing or after snR30 action (see discussion). Consistent with the hypothesis that yRrp7 might associate with the 18S sequence in the absence of snR30, the top snoRNA in the yRrp7 CRAC data set was indeed snR10, and not snR30 (21), and we were unable to observe a direct interaction between recombinant GST-yRrp7 and *in vitro* transcribed snR30 (Figure 3D, upper panel), or, in fact, snR10 (data not shown). Curiously, GST-yRrp7 strongly interacted with U17 (Figure 3D, lower panel), but yRrp7 had also previously been reported to interact with another, functionally unrelated box H/ACA snoRNA, snR5, *in vitro* (21). It is therefore unclear if the observed yRrp7/U17 interaction is biologically relevant. Given that the U17 and snR5 snoRNAs both lack an extended internal hairpin (Supplementary Figure S4D and data not shown), it is also possible that the snR30-specific internal hairpin might interfere with yRrp7 binding to snR30 in our *in vitro* assay.

We further assessed whether yUtp23 directly associates with other known snR30 binding factors. We detected a reproducible interaction between yUtp23 and the nucleolar protein yKri1 (Figure 3, panel C), which, like yUtp23, requires snR30 for pre-ribosomal recruitment (15). Furthermore, we also found that yUtp23 directly interacted with yNhp2, one of the four core box H/ACA proteins, while no interaction was detected with the U3-specific yRrp9 protein. It is likely that the observed protein–protein interaction between yUtp23 and yNhp2 contributes to yUtp23 binding to the snR30 snoRNP. This might explain why the internal hairpin in snR30, which is important for binding of yUtp23 to snR30 *in vitro* (Figure 2F), is not essential for survival *in vivo* (18).

Finally, we also tested human homologues of key ES6-binding factors (hUTP24, hROK1/DDX52 and hRRP7)

as well as the core box H/ACA protein hNHP2, for their ability to interact with recombinant GST-tagged human hUTP23 (panels E and F, respectively). Importantly, positive interactions detected in yeast (panels B and C) were also detected in the human system. Human hUTP23 interacted with hUTP24, hRRP7, hNHP2 and hROK1/DDX52. This indicates that key protein–protein interactions at the 18S ES6 rRNA region are evolutionarily conserved.

In conclusion, the observed direct interactions between yUtp23/hUTP23 and factors that also bind to the ES6 region or the snR30/U17 snoRNAs, in combination with the positioning of their crosslink sites on the individual RNAs, provide important insights into their spatial and temporal association within the yeast and human SSU processome (see discussion).

An intact PIN domain and Zinc finger in human hUTP23 are both required for 18S rRNA maturation

In budding yeast, yUtp23 contains only two (D31 and D123) of the four possible acidic residues that make up the characteristic PIN domain endonuclease catalytic center (Figure 4A and Supplementary Figure S6), and mutational analyses of these two residues strongly suggested a non-enzymatic role for yUtp23 within the yeast SSU processome (4,14). The PIN domain of human hUTP23, on the other hand, contains conserved acidic residues at three positions (D31, E68 and D122) and it was shown for the PIN domain of the NMD endonuclease SMG6, that a triad of acidic residues is sufficient for catalytic activity (33). Both human and yeast Utp23 proteins also share a conserved CCHC Zinc finger motif with a predicted role in nucleic acid binding. Single point mutations in the Zinc finger of yeast yUtp23 are lethal (14), but the essential role of this motif has not yet been determined.

RNAi-mediated knockdown of human hUTP23 results in 30S pre-rRNA accumulation, indicating that the presence of hUTP23 is required for the A0, 1 and 2a pre-rRNA cleavages (12) (Figure 4B). The PIN endonuclease hUTP24 is responsible for cleavages at sites 1 and 2a (6,7), but the identity of the enzyme that cleaves at site A0 remains unknown. To investigate a potential catalytic function of human hUTP23 in site A0 cleavage, and to understand the role of its conserved Zinc finger motif, we established an RNAi-rescue system for hUTP23 in HEK293 cells. For this, cells were stably transfected with an empty control vector (pcDNA5) or plasmids encoding C-terminally 2xHA-tagged hUTP23 carrying silent mutations in the hUTP23 open reading frame rendering it resistant to RNAi-knockdown when the endogenous hUTP23 mRNA is targeted. Cells expressing hUTP23 WT or mutant hUTP23 proteins with a single catalytic site mutation in the PIN domain (D31N) or a point mutation in the Zinc finger motif (C103A) were compared. Using a titratable *TET* promoter, conditions were established to express WT and mutant proteins at equivalent levels (Figure 4C). WT and mutant proteins showed some variation in their response to induction by tetracycline and it was therefore difficult to titrate the expression of all HA-tagged proteins to a level equivalent to that of the endogenous protein. To address this, two separate WT samples are presented in Figure 4C,

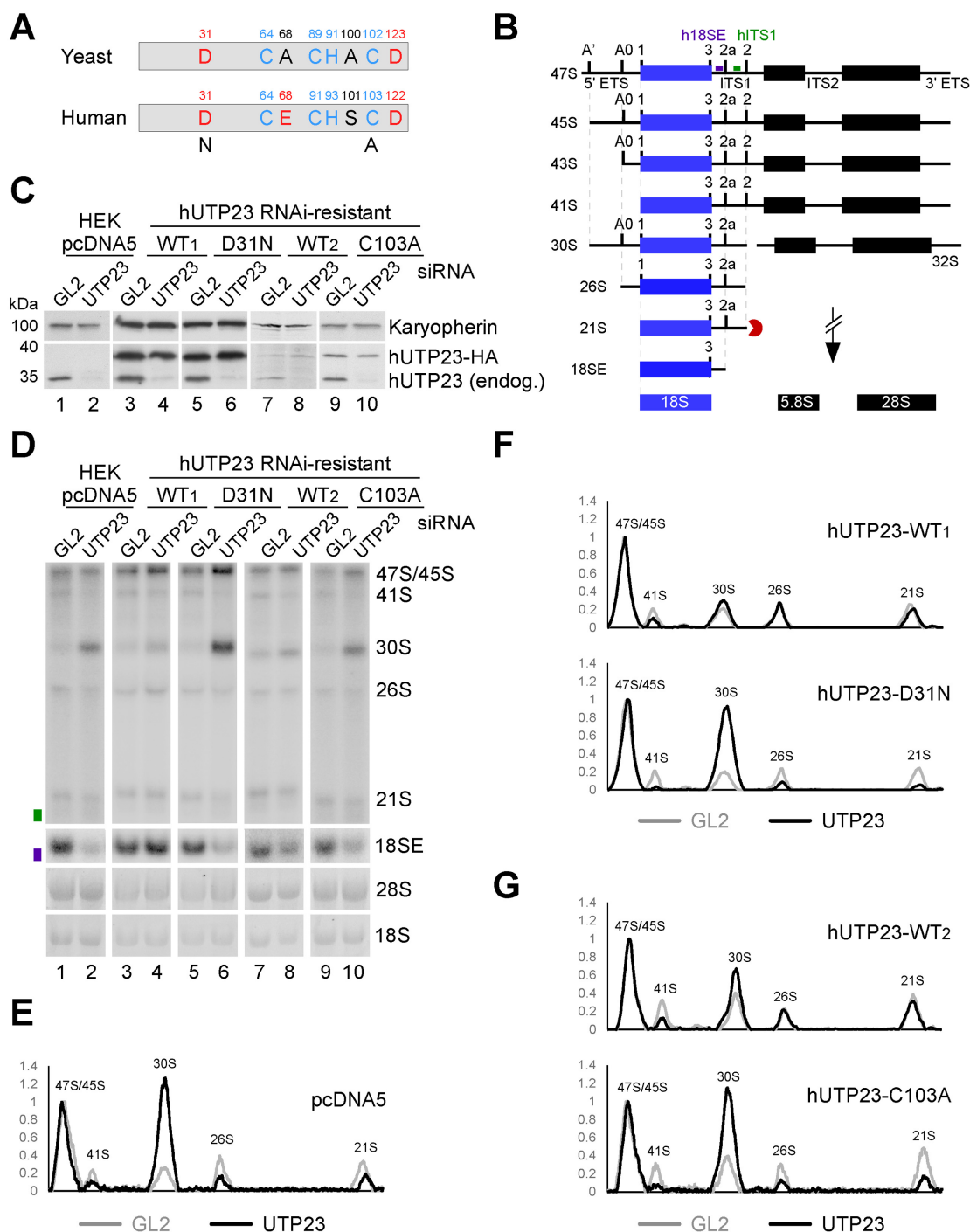


Figure 4. Intact PIN domain and Zinc finger motifs in human hUTP23 are essential for 18S maturation. (A) Cartoon depiction of the PIN domain and CCHC Zinc finger motifs in yeast yUtp23 and human hUTP23. Acidic and non-acidic residues in the proposed catalytic centre of the PIN domain are marked in red and black, respectively. Blue: conserved Zinc finger residues. Point mutations generated in the PIN domain (D31N) or the Zinc finger (C103A) of hUTP23 are indicated. (B) Schematic representation of the key steps in human ribosome biogenesis. Radiolabeled probes used for northern blotting (h18SE; purple and hITS1; green) are marked above the 47S precursor. ETS: external transcribed spacer; ITS: internal transcribed spacer. (C) HEK293 cells were stably transfected with a control plasmid (pcDNA5) or constructs encoding wild type (WT) or mutant forms of HA-tagged hUTP23 (D31N, C103A) and treated with tetracycline to induce protein expression. The hUTP23 coding sequence was modified to render the expressed mRNA resistant to the hUTP23 siRNA. Protein extracted from control cells (GL2), or those depleted of endogenous hUTP23 (UTP23) by RNAi, was separated by SDS-PAGE and transferred to nitrocellulose membrane. Protein levels were analyzed by immunoblotting using antibodies specific for hUTP23 (lower panel) or Karyopherin (upper panel) as loading control. (D) RNA from stably transfected and RNAi-treated HEK293 cells as shown in panel C was analyzed by northern blotting using probes hybridizing to the 5' end of ITS1 (h18SE, purple rectangle) or downstream of 2a (hITS1, green rectangle). Pre-rRNAs were detected using a PhosphorImager and total rRNA (28S/18S) was visualized by ethidium bromide (EtBr) staining. RNA species are labeled on the right. (E–G) RNA levels from panel D were normalized to the 47S/45S pre-rRNAs and plotted for each GL2 (gray) or hUTP23 (black) knockdown. The identity of each peak is indicated.

with expression levels that are directly comparable to either the D31N (WT1, compare lanes 3–6) or the C103A (WT2, compare lanes 7–10) mutant, respectively. Cells were transfected with either a siRNA specifically targeting endogenous hUTP23 or a control siRNA targeting firefly luciferase (GL2) (34). After siRNA treatment for 72 h, the expression of endogenous and HA-tagged hUTP23 proteins was analyzed by immunoblotting. The hUTP23-specific siRNA significantly reduced endogenous protein levels, whereas the RNAi-resistant HA-tagged hUTP23 proteins were unaffected (Figure 4C).

Total RNA was extracted from RNAi-treated cells and pre-rRNA processing was analyzed by Northern hybridization using probes complementary to the 5' end of ITS1 ('h18SE') or downstream of the 2a cleavage site ('hITS1') (Figure 4, panels B and D–G). Depletion of hUTP23 resulted in a significant accumulation of the 30S pre-rRNA, indicative of strongly reduced early cleavages at the A0, 1 and 2a sites, and consequently, a substantial reduction in 18SE levels as previously reported (Figure 4, panels D and E) (12). A 5'-extended 30S precursor (referred to as 30SL5' (12) or 34S (35)), did not accumulate, suggesting that cleavage at A' was not affected (Figure 4D, compare lanes 1 and 2). In cells expressing moderately high levels of wild type HA-tagged hUTP23 (WT1), knockdown of endogenous hUTP23 had no effect on either 18SE or 30S pre-rRNA levels (compare lanes 3 and 4, see panel F for quantification), demonstrating that HA-tagged hUTP23 can functionally replace the endogenous protein in our system. While expression of lower levels of the WT protein (WT2; note that levels are slightly lower than the endogenous protein) did not rescue the processing phenotype to the same extent, a significant reduction of the 30S levels compared to the hUTP23 depletion phenotype was still observed (compare lanes 8 and 2, see panels G and E for quantification).

Surprisingly, expression of either the PIN (D31N, Figure 4D, lane 6) or the Zinc finger (C103A, lane 10) hUTP23 mutants, after hUTP23 knockdown, resulted in strong 30S accumulation compared to their respective WT control (see panels F and G for quantification), a phenotype identical to hUTP23 depletion in the absence of exogenous protein (lane 2). This indicates that the putative active site in the PIN domain of hUTP23 and the integrity of the conserved Zinc finger are both important for hUTP23 function in pre-rRNA processing.

We next investigated the reason(s) why the PIN and Zinc finger mutations impeded hUTP23 function. The significant accumulation of the 30S pre-rRNA caused by expression of the hUTP23 D31N PIN mutant *in vivo* (Figure 4D, lane 6) would be readily explained by a loss of enzymatic activity at site A0. Since site A0, 1 and 2a cleavages are linked, a defect in A0 cleavage would also lead to defects in the cleavages at sites 1 and 2a, which is fitting with the substantial increase in 30S pre-rRNA levels seen upon expression of the hUTP23 D31N PIN mutant. We therefore tested whether recombinant hUTP23 (see Supplementary Figure S5A) would cleave an RNA substrate containing human site A0. However, attempts to demonstrate nuclease activity were so far unsuccessful (data not shown and see discussion).

Immunofluorescence experiments using cells expressing HA-tagged proteins demonstrated that, in contrast to fibrillarin, which is exclusively found in the dense fibrillar component (29), wild-type hUTP23 is localized throughout the nucleolus (Figure 5A). Similar to the wild type protein, the mutant hUTP23 proteins were also found throughout the nucleolus. However, immunoprecipitation experiments revealed that only the WT hUTP23 protein was stably associated with the human homologue of snR30, the U17 snoRNA (Figure 5B, lane 4). The PIN mutant exhibited considerably weaker, but still detectable association with the U17 snoRNA (see lane 6), whereas the C103A Zinc finger mutant did not co-precipitate the U17 snoRNA above background (lane 8). Consistent with our yeast data (Figure 2), recombinant WT hUTP23 directly bound to *in vitro* transcribed U17 snoRNA (Figure 5C and 5D, upper panels). The D31N PIN mutant associated with U17 with similar affinity (Figure 5C, lower panel), whereas the C103A Zinc finger mutant exhibited significantly reduced binding (Figure 5D, lower panel). Interestingly, the same result was observed when we analyzed the binding of WT recombinant yeast yUtp23 and a C102A Zinc finger mutant (Supplementary Figure S5B) to either yeast snR30 or human U17 (Figure 5E). The WT yUtp23 protein efficiently bound both U17 and snR30, but in contrast, the C102A Zinc finger mutant showed significantly reduced binding to both snoRNAs at low protein concentration. In the presence of high levels of the C102A mutant, about half of the U17 substrate was still bound compared to the wild type protein, whereas resolvable complexes of bound snR30 or free RNA could not be detected. This result suggests non-stoichiometric and likely non-specific protein-snR30 interactions when the mutant protein is present at high concentration in the assay. The *in vitro* RNA binding results indicate that the severe growth defect in yeast (14) and the pre-rRNA processing defect in humans (Figure 4), seen upon expression of the yUtp23/hUTP23 Zinc finger mutants, are likely due to a defect in yUtp23/hUTP23 binding to the snR30 and U17 snoRNAs, respectively.

Taken together, our data indicate that an intact PIN domain in hUTP23 and efficient binding of hUTP23 to the U17 snoRNP are both required for 18S maturation in humans.

DISCUSSION

Here, we present data suggesting that the yeast yUtp23-snR30 and human hUTP23-U17 complexes may act as hubs to coordinate the binding and release of factors at the 18S ES6 region. We also show that the PIN domain active site amino acid, D31, is needed for pre-rRNA processing suggesting that hUTP23 may actually be an endonuclease.

Using *in vivo* crosslinking in actively growing yeast cells, we have identified snR30 and the 18S ES6 region as *bona fide* RNA-binding targets for yUtp23 (Figures 1 and 2, and Supplementary Figures S2, S3 and S4). Direct protein-RNA interactions were validated *in vitro* using recombinant proteins and *in vitro* transcribed RNA and we also showed that stable interactions with snR30/U17 required an intact Zinc finger in yUtp23/hUTP23 (Figure 5). The top of the 3' hairpin in snR30 was predicted to bind a pro-

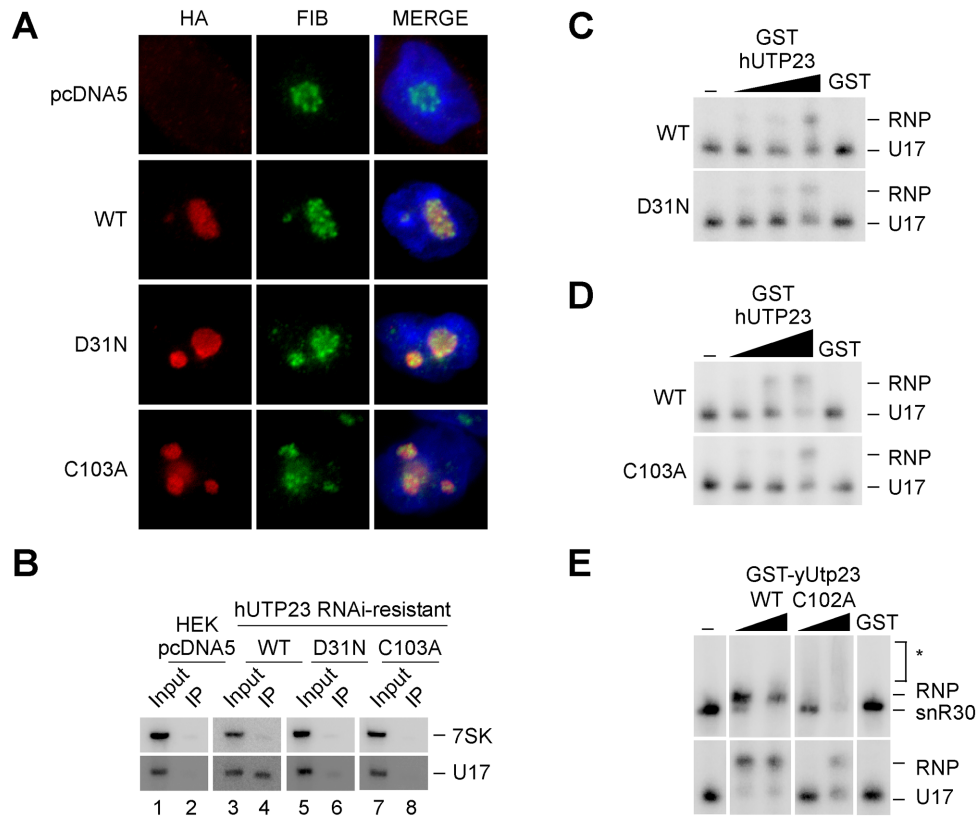


Figure 5. The hUTP23 Zinc finger mutant exhibits typical nucleolar localization, but impaired *in vivo* and *in vitro* binding to U17. (A) Cells expressing WT or mutant forms of HA-tagged hUTP23 (D31N, C103A) or no HA-tagged protein (pcDNA5) were harvested and processed for immunofluorescence microscopy using an anti-HA antibody (red). Cells were counterstained with an antibody against endogenous fibrillarin (FIB, green) as a nucleolar marker and DAPI (blue) to highlight the nucleus. From left to right: Immunofluorescence signals for the anti-HA and anti-fibrillarin antibodies and the merged image. (B) Soluble lysates from cells, as shown in panel A, were subjected to immunoprecipitation using an anti-HA antibody. Co-precipitated RNA was extracted and separated by denaturing acrylamide electrophoresis, analyzed by northern blotting using radiolabeled probes against the U17 snoRNA and 7SK as a loading control and visualized using a PhosphorImager. A total of 1.25% of the input material was loaded. (C) EMSA showing the binding of recombinant wild type (WT, top panel) or D31N PIN mutant (bottom panel) human GST-hUTP23 proteins (450, 900 and 1800 nM) or GST (1800 nM) to *in vitro* transcribed radiolabeled U17 snoRNA. RNA and RNP complexes were analyzed as in Figure 2B. (D) Binding of WT (top panel) or the C103A Zn finger mutant (bottom panel) human GST-hUTP23 proteins (750, 1500 and 3000 nM) or free GST (3000 nM) to the U17 snoRNA. Analysis as described in panel C. (E) Binding of recombinant WT and Zinc finger mutant (C102A) yeast GST-yUtp23 proteins (250 and 2500 nM) or GST (2500 nM) to snR30 (top) or U17 (bottom) RNAs. Asterisks: unresolved C102A mutant-snR30 complexes, likely due to non-stable protein-RNA interactions. Analysis as described in panel C.

tein essential for snR30 function in 18S rRNA processing (19) and yUtp23 has been identified, alongside yKri1, as a novel non-core snoRNP protein co-purifying with snR30 (15). Interestingly, we detected a direct protein-protein interaction between yUtp23 and yKri1 (Figure 3C). This result was surprising, given that these proteins have been suggested to associate with snR30 in a mutually exclusive manner (15). It was also shown in the same study that yUtp23, but not yKri1, is essential for snR30 release. The observed direct interaction between both proteins may therefore suggest that they are involved in distinct, but possibly consecutive steps of snR30 function, with yUtp23 playing a more active role in snR30 release. Our *in vivo* crosslinking and *in vitro* RNA binding studies have revealed that yUtp23 primarily contacts the top part of the 3' hairpin and the internal hairpin of snR30 (Figures 2 and 6A). Our data further showed that all three hairpins in snR30 are important for efficient binding of yUtp23 *in vitro*. The finding that the internal hairpin is a major snR30 binding site of the

essential yUtp23 protein *in vivo* and *in vitro* is surprising given that this region of the snoRNA has been shown to be dispensable for snR30 function *in vivo* (18). Interestingly, we also discovered an evolutionarily conserved interaction between yUtp23/hUTP23 and the H/ACA snoRNP core component, yNhp2/hNHP2 (Figure 3) and yeast yUtp23 was previously reported to interact with another H/ACA snoRNP component, yGar1 (36). We believe that the interactions between yUtp23 and both yNhp2 and yGar1, which are expected to also bind to the upper part of the snR30 3' hairpin (Figure 6A), are sufficient to tether yUtp23 on the snR30 snoRNP and thereby compensate for the lack of the internal hairpin *in vivo*. While further experiments are clearly needed to clarify the exact role of yKri1 with respect to snR30 function, our data strongly suggest that yUtp23 is one of the predicted essential snR30-specific protein(s) that bind the 3' hairpin in the snoRNA.

In vivo protein-RNA crosslinks were also found between yUtp23 and the ES6 region of the 18S rRNA. The speci-

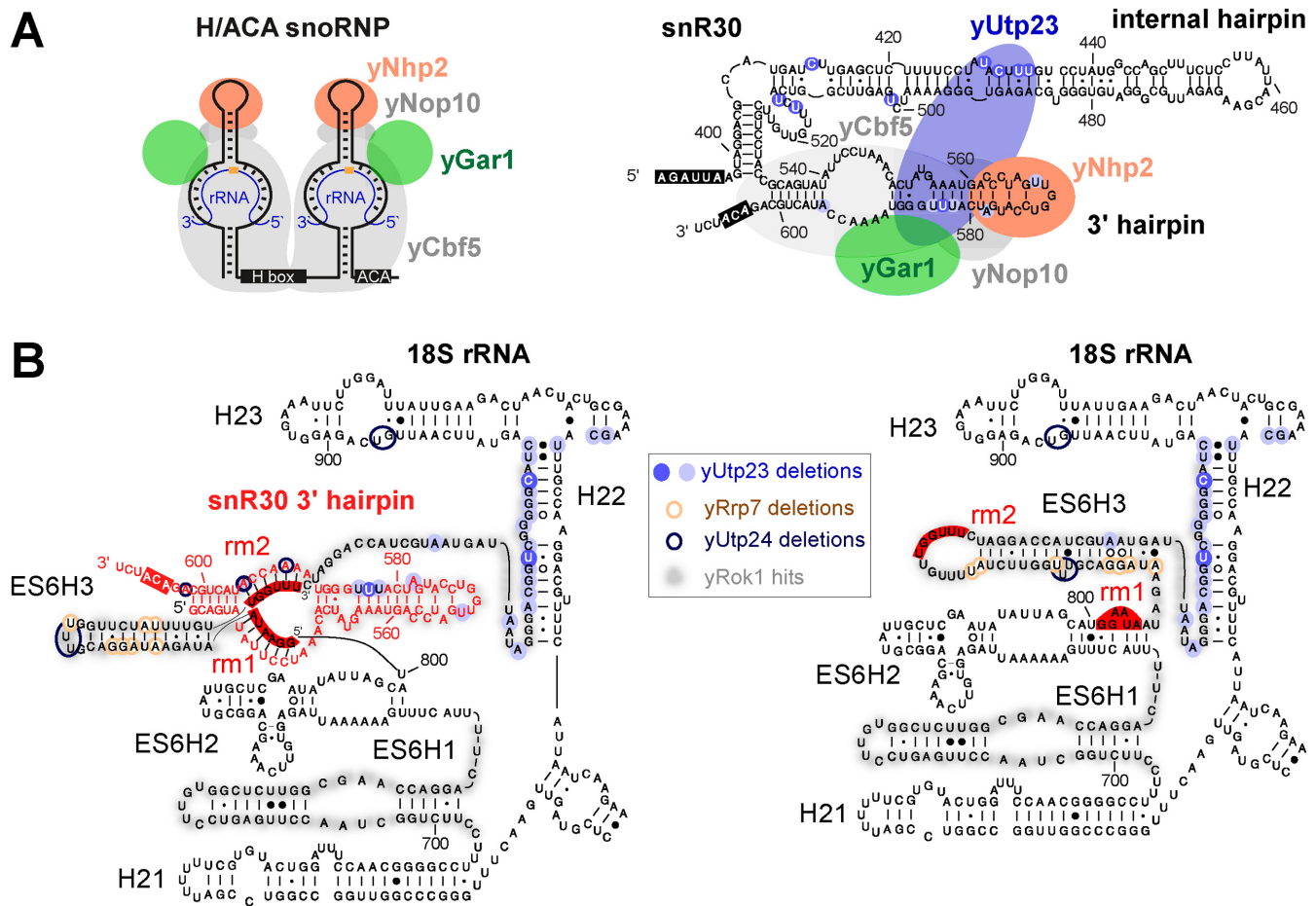


Figure 6. Yeast yUtp23 is central to the integration of ES6 binding factors and snR30 release. (A) Left panel: Cartoon of a typical modification H/ACA box yeast snoRNP depicting the positions of the core H/ACA snoRNP proteins yCbf5, yGar1, yNhp2, and yNop10. H and ACA boxes are shown in black. Orange circle: target nucleotide for pseudouridylation in the rRNA (blue). Right panel: Cartoon of the 3' terminal portion of yeast snR30 with the predicted positions of yUtp23 and core H/ACA snoRNP proteins. yUtp23 microdeletion sites in the internal hairpin and 3' hairpin are represented as blue circles, shades represent peak height as in Figure 2C. The H and ACA boxes are shown in black. (B) Predicted structure of the yeast 18S ES6 region during snR30 base-pairing (left panel) and in the mature rRNA after snR30 release (right panel). Precise crosslinking sites of yUtp23 (blue circles, shades indicate peak height as in Figures 1 and 2), yUtp24 (dark blue) and yRrp7 (brown), and yRok1 crosslinking regions (shaded in grey) on the snR30 3' hairpin (red) and 18S rRNA sequences (black). Binding sites for snR30 in the ES6 region (rm1 and rm2) and the ACA box are highlighted in red.

ficity of this interaction was validated with an *in vitro* RNA binding experiment (Supplementary Figure S2D). Primer extension analyses of RNAs crosslinked to both yeast yUtp23 and human hUTP23 proteins further revealed that this interaction is evolutionarily conserved (Supplementary Figure S3). Our data therefore show that yUtp23/hUTP23 binds directly to two different RNAs, snR30/U17 and the 18S rRNA ES6 region. Interestingly, the main 18S rRNA interaction site for yUtp23/hUTP23 is in close proximity to the snR30/U17 base-pairing regions in ES6 (rm1 and rm2) and near the known crosslinking sites for yeast yRok1, yRrp7, yDhr1, yUtp24 and yRrp5 (Figure 6B). Furthermore, we could show direct protein–protein interactions between yUtp23 and yRok1, yRrp7 and yUtp24 and their human equivalents hUTP23, hRok1, hRRP7 and hUTP24 (Figure 3). The yUtp23–snR30 and hUTP23–U17 complexes therefore appear to form central hubs within the yeast and human SSU processomes that potentially coordinate the binding of factors to, and their subsequent release from, ES6. Indeed, it has already been shown that

yUtp23 is needed for snR30 release from the pre-rRNA (15), a process involving the RNA helicase yRok1 and the compaction factor yRrp5 (23,24).

In vivo crosslinking studies in yeast revealed that several of the protein contact sites in ES6 overlap (Figure 6B), suggesting sequential and coordinated protein binding. Consistent with this idea, some of the ES6 binding proteins, such as yRrp5, yRrp7 and yUtp24, are still present in the 90S pre-ribosome after snR30 and yUtp23 have left the complex (37). Notably, snR30 is required for stable association of yRrp7 with pre-ribosomes, but not vice versa (21). Given the dramatic change in the structure of the yRrp7 binding site in ES6 seen before and after snR30 release (Figure 6B, compare left and right panels), it is unlikely that yRrp7 binds to both structures. yRrp7 is present in 90S pre-ribosomes lacking snR30 and, within this complex, probably still bound to the pre-rRNA after snR30 release. We therefore hypothesize that yRrp7 recognizes the post-snR30 ES6 structure. Supporting this idea, yRrp7 mainly crosslinked to snR10 instead of snR30

(21) and we did not detect stable *in vitro* interactions between yRrp7 and snR30 (Figure 3D). yUtp23 may therefore, together with snR30, function as a platform to recruit yRrp7 to the pre-ribosome through protein–protein interactions. Subsequently, the yRrp7 binding site may then be generated by the yUtp23-dependent release of snR30 and the re-structuring of ES6. Protein–protein interactions between yUtp23 and yRok1 suggest a similar scenario, where yUtp23 could contribute to the role of yRok1 in releasing snR30 from the pre-ribosome. In this context, we were surprised to observe an *in vitro* interaction between yeast yRrp7 and human U17 (Figure 3D). However, future studies in human cells will be needed to clarify whether hRRP7 and U17 interact *in vivo* and whether this is biologically relevant.

Using an RNAi-rescue system, we have also analyzed the function of human hUTP23. Knockdown of hUTP23 inhibits cleavages at sites A0, 1 and 2a (12). These defects were rescued by re-expression of the WT protein, but not hUTP23 containing point mutations in either the Zinc finger or the PIN domain (Figure 4), although WT and both mutant proteins localized to the nucleolus as expected (Figure 5). We observed that the hUTP23 Zinc finger mutant (C103A) showed strongly reduced association with the U17 snoRNA *in vitro* and *in vivo*. The equivalent point mutation in yeast yUtp23 (C102A) also impaired *in vitro* interactions with snR30 (Figure 5) and caused lethality (14). In yeast, snR30 is needed for correct positioning of yUtp23 within the SSU processome to enable yUtp23 binding to the 35S pre-rRNA (15). Our comparative analyses therefore indicate that a direct protein–RNA interaction between yUtp23/hUTP23 and snR30/U17 is evolutionarily conserved and is dependent on the integrity of the Zinc finger.

An intact PIN domain, on the other hand, is not essential for yUtp23 function in yeast (4,14). In humans, however, the PIN domain mutant (D31N) did not rescue hUTP23 function despite correct nucleolar localization. Binding of the recombinant mutant protein to the U17 snoRNA was not affected *in vitro*, indicating that there is no major defect in protein folding caused by the single point mutation in the PIN domain. We did, however, notice significantly decreased co-precipitation of U17 with the hUTP23 PIN mutant *in vivo*. It is therefore possible that the D31N mutation might interfere with other essential hUTP23 interactions within the human pre-40S particle that are needed for U17 association. In yeast and other fungi, two of the four PIN domain active site amino acids are present in yUtp23 (Supplementary Figure S6), however, in metazoan and plant UTP23, three of the key active-site amino acids are conserved. In some PIN-domain proteins, such as SMG6, this is sufficient for nuclease activity (33). Our attempts to detect nuclease activity in human hUTP23, using a variety of human pre-ribosomal RNA substrates containing site A0 have so far been unsuccessful. While the current lack of *in vitro* cleavage data could indicate a non-catalytic role for hUTP23 in human ribosome biogenesis, the strong pre-rRNA processing defects observed upon expression of the PIN mutant *in vivo* make it likely that instead the right *in vitro* substrate has not yet been identified. In addition, other protein factors may be needed for correct RNA folding *in vitro* and/or activation of

hUTP23. Indeed, the human PIN domain protein hNOB1 was recently shown to require a co-factor, hCINAP1, for cleavage at site 3 *in vitro* (11). Future experiments will reveal whether hUTP23 is an active enzyme and whether a co-factor(s) is required for its activity.

ACCESSION NUMBERS

GEO database (<http://www.ncbi.nlm.nih.gov/geo/>): identifier GSE87238.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

The authors thank David Tollervey (Wellcome Trust Centre for Cell Biology, University of Edinburgh) for supporting the early part of this work. The authors also thank Grzegorz Kudla (MRC Institute of Genetics and Molecular Medicine, University of Edinburgh) for help with the CRAC analyses.

FUNDING

Royal Society [UF100666 and RG110357 to C.S.]; Wellcome Trust [092076 to N.J.W.]; BBSRC project [BB/F006853/1 to N.J.W.]; ERASMUS internship [to F.W.]. Funding for open access charge: BBSRC [BB/F006853/1 to N.J.W.].

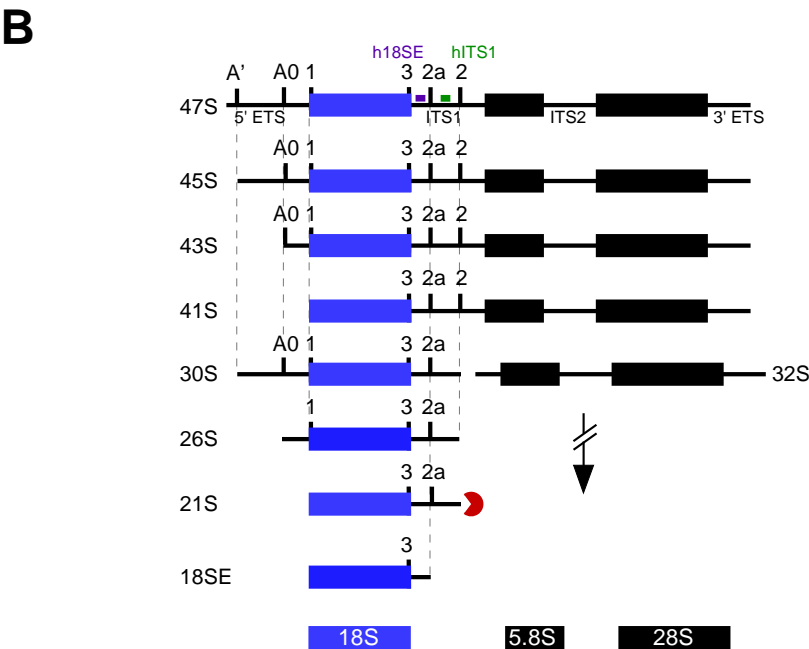
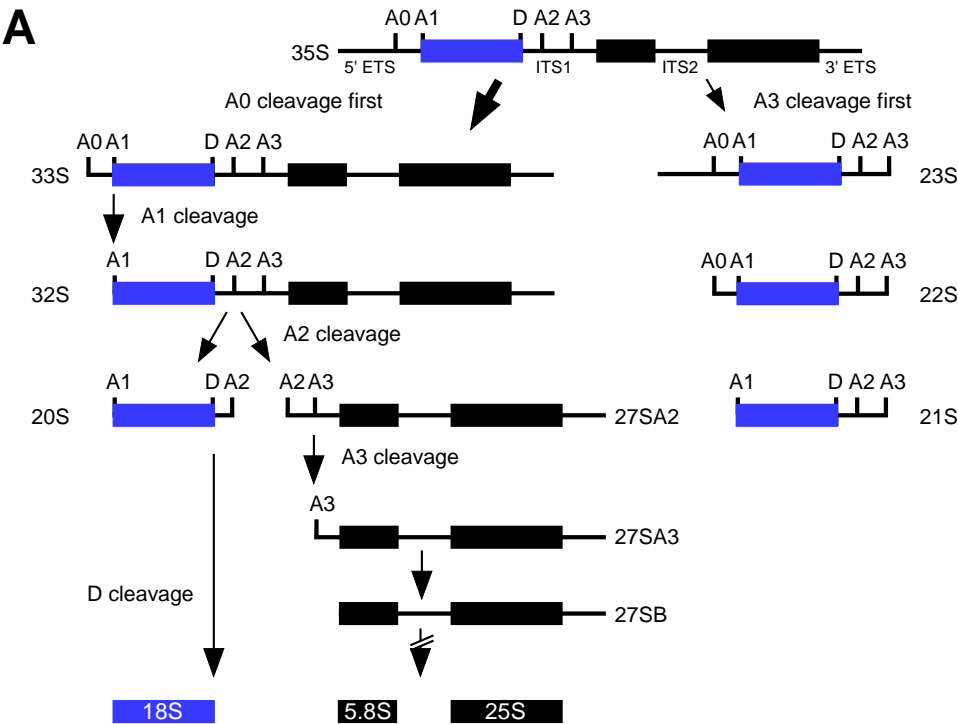
Conflict of interest statement. None declared.

REFERENCES

- Henras,A.K., Plisson-Chastang,C., O'Donohue,M.F., Chakraborty,A. and Gleizes,P.E. (2015) An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA*, **6**, 225–242.
- Phipps,K.R., Charette,J. and Baserga,S.J. (2011) The small subunit processome in ribosome biogenesis-progress and prospects. *Wiley Interdiscip. Rev. RNA*, **2**, 1–21.
- Watkins,N.J. and Bohnsack,M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA*, **3**, 397–414.
- Bleichert,F., Granneman,S., Osheim,Y.N., Beyer,A.L. and Baserga,S.J. (2006) The PINc domain protein Utp24, a putative nuclease, is required for the early cleavage steps in 18S rRNA maturation. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 9464–9469.
- Sloan,K.E., Mattijssen,S., Lebaron,S., Tollervey,D., Puijn,G.J. and Watkins,N.J. (2013) Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *J. Cell Biol.*, **200**, 577–588.
- Tomecki,R., Labno,A., Drazkowska,K., Cysewski,D. and Dziembowski,A. (2015) hUTP24 is essential for processing of the human rRNA precursor at site A1, but not at site A0. *RNA Biol.*, **12**, 1010–1029.
- Wells,G.R., Weichmann,F., Colvin,D., Sloan,K.E., Kudla,G., Tollervey,D., Watkins,N.J. and Schneider,C. (2016) The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. *Nucleic Acids Res.*, **44**, 5399–5409.
- Fatica,A., Oeffinger,M., Dlakic,M. and Tollervey,D. (2003) Nob1p is required for cleavage of the 3' end of 18S rRNA. *Mol. Cell. Biol.*, **23**, 1798–1807.

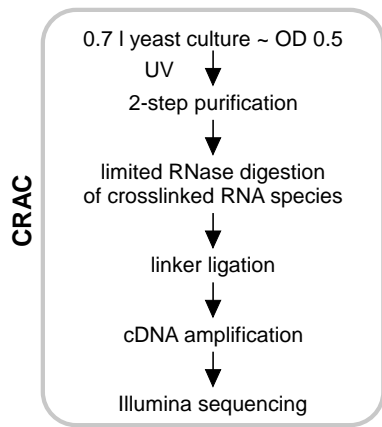
9. Pertschy, B., Schneider, C., Gnadig, M., Schafer, T., Tollervey, D. and Hurt, E. (2009) RNA helicase Prp43 and its co-factor Pfa1 promote 20 to 18 S rRNA processing catalyzed by the endonuclease Nob1. *J. Biol. Chem.*, **284**, 35079–35091.
10. Lebaron, S., Schneider, C., van Nues, R.W., Swiatkowska, A., Walsh, D., Botcher, B., Granneman, S., Watkins, N.J. and Tollervey, D. (2012) Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nat. Struct. Mol. Biol.*, **19**, 744–753.
11. Bai, D., Zhang, J., Li, T., Hang, R., Liu, Y., Tian, Y., Huang, D., Qu, L., Cao, X., Ji, J. *et al.* (2016) The ATPase hCINAP regulates 18S rRNA processing and is essential for embryogenesis and tumour growth. *Nat. Commun.*, **7**, 12310.
12. Sloan, K.E., Bohnsack, M.T., Schneider, C. and Watkins, N.J. (2014) The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA*, **20**, 540–550.
13. Wang, M., Anikin, L. and Pestov, D.G. (2014) Two orthogonal cleavages separate subunit RNAs in mouse ribosome biogenesis. *Nucleic Acids Res.*, **42**, 11180–11191.
14. Lu, J., Sun, M. and Ye, K. (2013) Structural and functional analysis of Utp23, a yeast ribosome synthesis factor with degenerate PIN domain. *RNA*, **19**, 1815–1824.
15. Hoareau-Aveilla, C., Fayet-Lebaron, E., Jady, B.E., Henras, A.K. and Kiss, T. (2012) Utp23p is required for dissociation of snR30 small nucleolar RNP from preribosomal particles. *Nucleic Acids Res.*, **40**, 3641–3652.
16. Morrissey, J.P. and Tollervey, D. (1993) Yeast snR30 is a small nucleolar RNA required for 18S rRNA synthesis. *Mol. Cell. Biol.*, **13**, 2469–2477.
17. Lemay, V., Hossain, A., Osheim, Y.N., Beyer, A.L. and Dragon, F. (2011) Identification of novel proteins associated with yeast snR30 small nucleolar RNA. *Nucleic Acids Res.*, **39**, 9659–9670.
18. Atzorn, V., Fragapane, P. and Kiss, T. (2004) U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. *Mol. Cell. Biol.*, **24**, 1769–1778.
19. Fayet-Lebaron, E., Atzorn, V., Henry, Y. and Kiss, T. (2009) 18S rRNA processing requires base pairings of snR30 H/ACA snoRNA to eukaryote-specific 18S sequences. *EMBO J.*, **28**, 1260–1270.
20. Martin, R., Hackert, P., Ruprecht, M., Simm, S., Bruning, L., Mirus, O., Sloan, K.E., Kudla, G., Schleiff, E. and Bohnsack, M.T. (2014) A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. *RNA*, **20**, 1173–1182.
21. Lin, J., Lu, J., Feng, Y., Sun, M. and Ye, K. (2013) An RNA-binding complex involved in ribosome biogenesis contains a protein with homology to tRNA CCA-adding enzyme. *PLoS Biol.*, **11**, e1001669.
22. Lebaron, S., Segerstolpe, A., French, S.L., Dudnakova, T., de Lima Alves, F., Granneman, S., Rappsilber, J., Beyer, A.L., Wieslander, L. and Tollervey, D. (2013) Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. *Mol. Cell*, **52**, 707–719.
23. Bohnsack, M.T., Kos, M. and Tollervey, D. (2008) Quantitative analysis of snoRNA association with pre-ribosomes and release of snR30 by Rok1 helicase. *EMBO Rep.*, **9**, 1230–1236.
24. Khoshnevis, S., Askenasy, I., Johnson, M.C., Dattolo, M.D., Young-Erdos, C.L., Stroupe, M.E. and Karbstein, K. (2016) The DEAD-box protein Rok1 orchestrates 40S and 60S ribosome assembly by promoting the release of Rrp5 from Pre-40S ribosomes to allow for 60S maturation. *PLoS Biol.*, **14**, e1002480.
25. Granneman, S., Petfalski, E. and Tollervey, D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J.*, **30**, 4006–4019.
26. Granneman, S., Kudla, G., Petfalski, E. and Tollervey, D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 9613–9618.
27. Wlotzka, W., Kudla, G., Granneman, S. and Tollervey, D. (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.*, **30**, 1790–1803.
28. Schneider, C., Anderson, J.T. and Tollervey, D. (2007) The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. *Mol. Cell*, **27**, 324–331.
29. Turner, A.J., Knox, A.A. and Watkins, N.J. (2012) Nucleolar disruption leads to the spatial separation of key 18S rRNA processing factors. *RNA Biol.*, **9**, 175–186.
30. Schneider, C., Kudla, G., Wlotzka, W., Tuck, A. and Tollervey, D. (2012) Transcriptome-wide analysis of exosome targets. *Mol. Cell*, **48**, 422–433.
31. Sardana, R., Liu, X., Granneman, S., Zhu, J., Gill, M., Papoulas, O., Marcotte, E.M., Tollervey, D., Correll, C.C. and Johnson, A.W. (2015) The DEAH-box helicase Dhr1 dissociates U3 from the pre-rRNA to promote formation of the central pseudoknot. *PLoS Biol.*, **13**, e1002083.
32. Lamanna, A.C. and Karbstein, K. (2009) Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 14259–14264.
33. Glavan, F., Behm-Ansmant, I., Izaurralde, E. and Conti, E. (2006) Structures of the PIN domains of SMG6 and SMG5 reveal a nuclease within the mRNA surveillance complex. *EMBO J.*, **25**, 5117–5125.
34. Elbashir, S.M., Harborth, J., Weber, K. and Tuschl, T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, **26**, 199–213.
35. Tafforeau, L., Zorbas, C., Langhendries, J.L., Mullineux, S.T., Stamatoopoulou, V., Mullier, R., Wacheul, L. and Lafontaine, D.L. (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol. Cell*, **51**, 539–551.
36. Tarassov, K., Messier, V., Landry, C.R., Radinovic, S., Serna Molina, M.M., Shames, I., Malitskaya, Y., Vogel, J., Bussey, H. and Michnick, S.W. (2008) An in vivo map of the yeast protein interactome. *Science*, **320**, 1465–1470.
37. Kornprobst, M., Turk, M., Kellner, N., Cheng, J., Flemming, D., Kos-Braun, I., Kos, M., Thoms, M., Berninghausen, O., Beckmann, R. *et al.* (2016) Architecture of the 90S Pre-ribosome: A Structural View on the Birth of the Eukaryotic Ribosome. *Cell*, **166**, 380–393.

Wells *et al.*, Figure S1

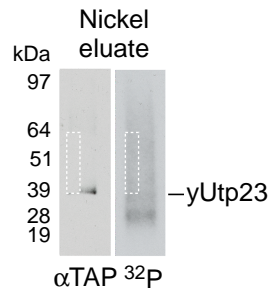


Wells *et al.*, Figure S2

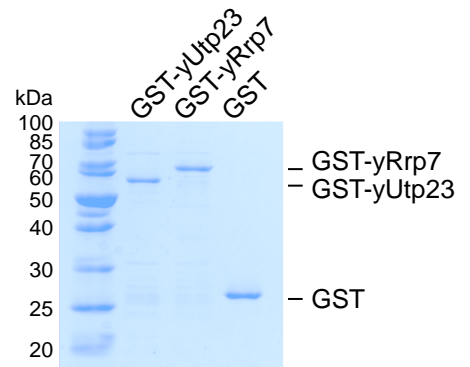
A



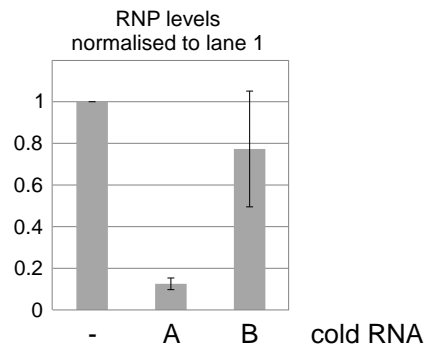
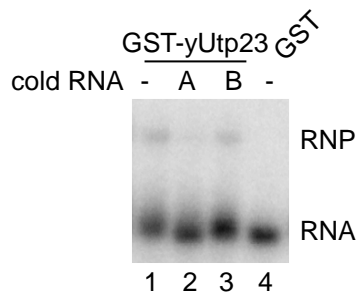
B



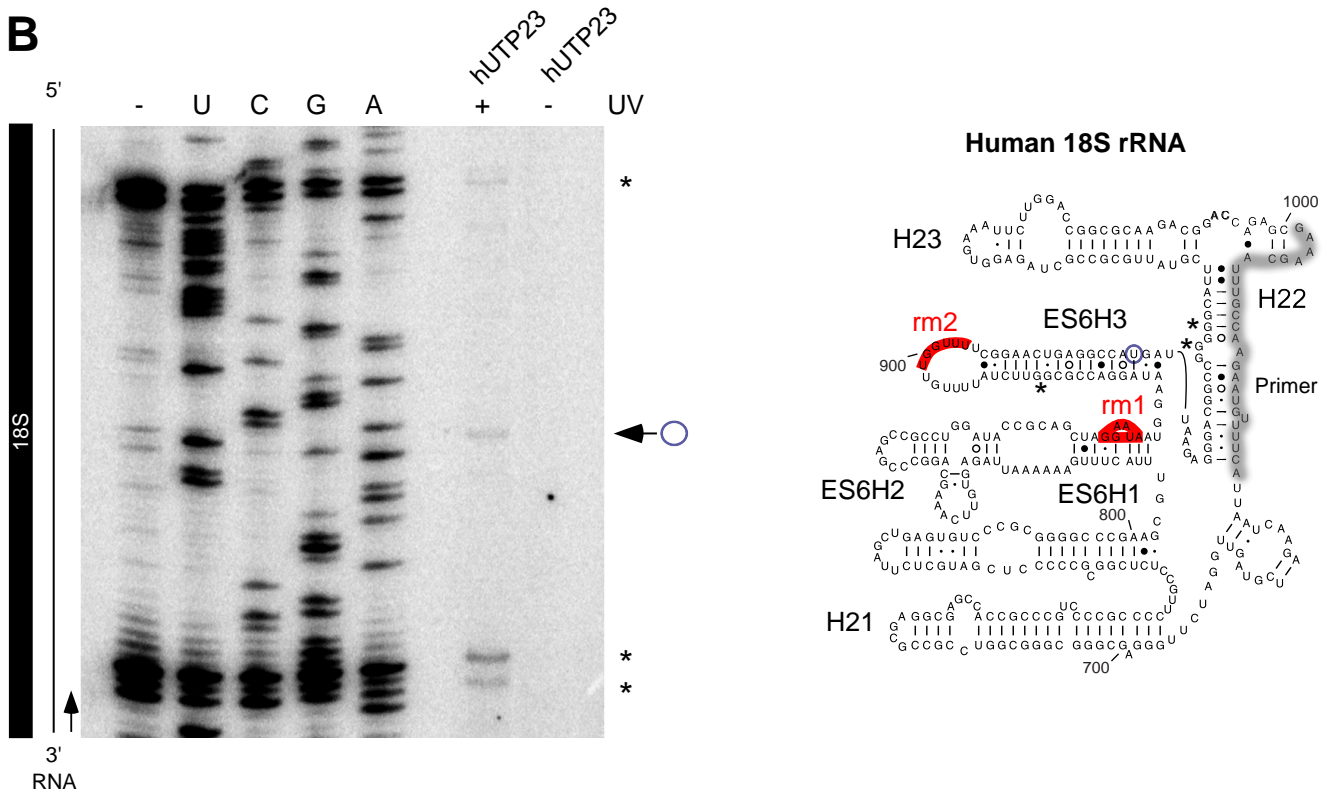
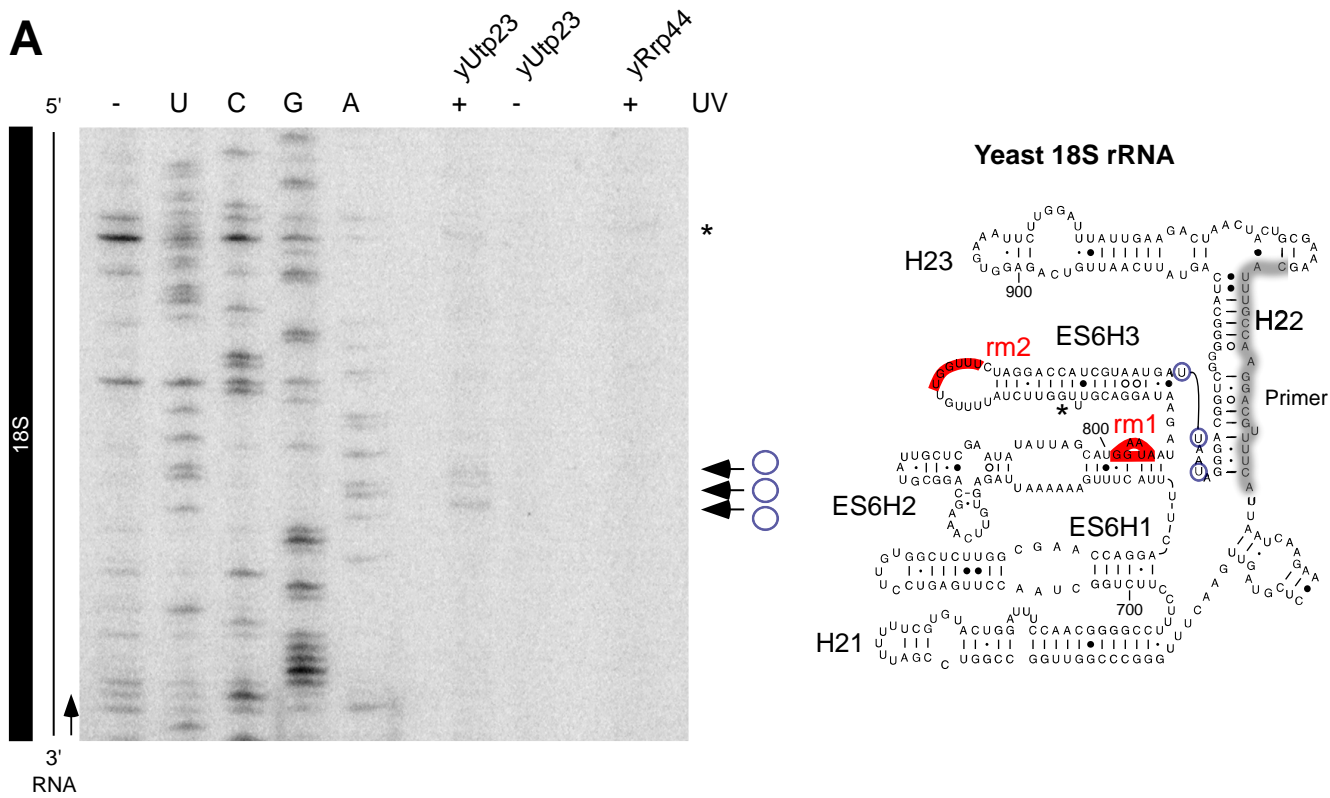
C



D



Wells *et al.*, Figure S3



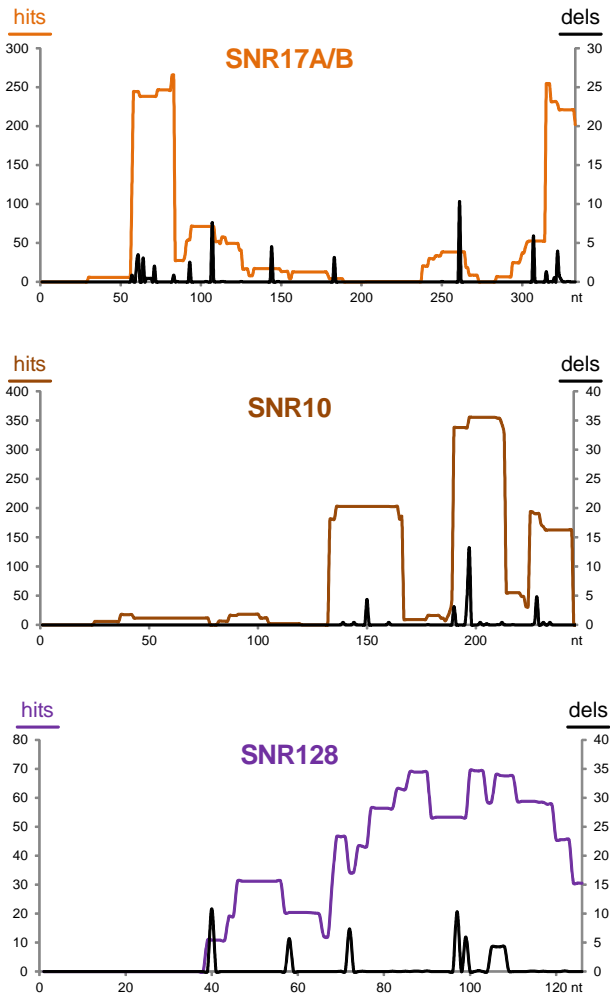
A

Type	Yeast	Human
H/ACA	snR30	U17
CD	snR17a/b/U3	U3
CD	snR128/U14	U14
H/ACA	snR10	NA

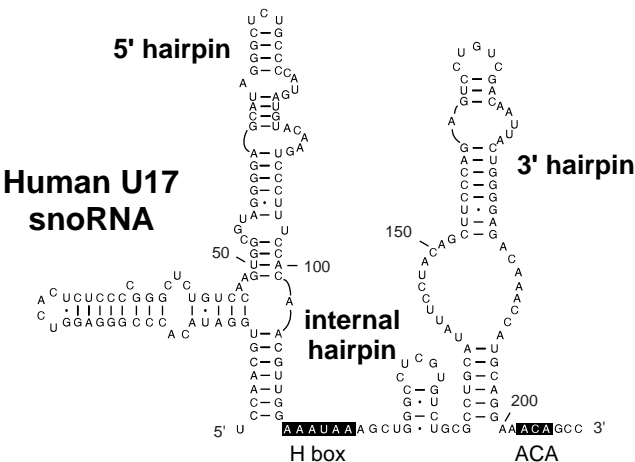
B

Top20 snoRNA (%)	Dataset #2	Dataset #4
SNR30	62.9	64.1
SNR17A/B	4.2	2.6
SNR190	3.9	1.5
SNR10	1.9	2.5
SNR128	1.9	0
SNR42	1.7	1.5
SNR86	1.7	1.9
SNR3	1.5	0.8
SNR4	1.2	0
SNR40	1.2	2.3
SNR82	1.1	2.2
SNR33	1	1
SNR31	0.9	0
SNR36	0.9	0
SNR11	0.8	0
SNR35	0.8	0
SNR37	0.8	1.4
SNR41	0.8	0.9
SNR84	0.8	1.5
SNR70	0.7	0.8
SNR13	0	1.7
SNR64	0	1
SNR66	0	1
SNR75	0	2.2
SNR77	0	2
SNR78	0	1.2
others	9.3	5.9

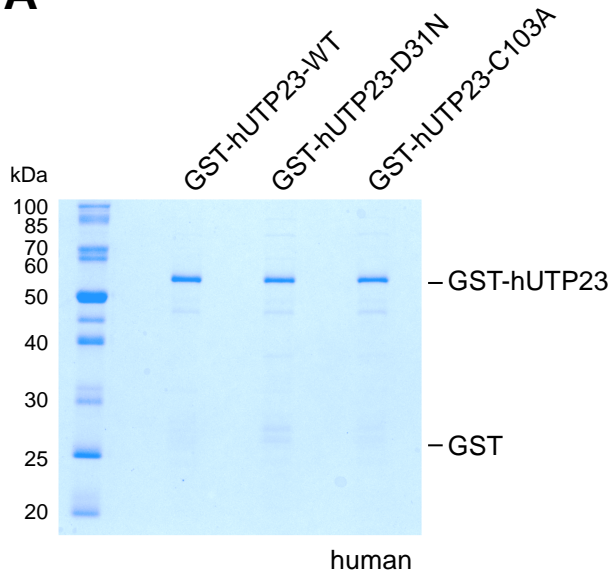
C



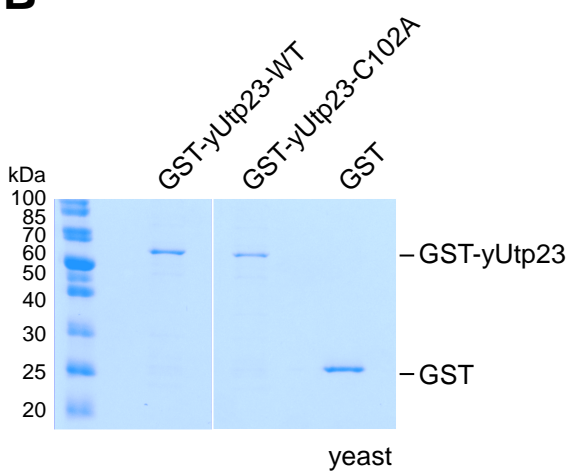
D



A



B



Wells et al., Figure S6

Utp23 CLUSTAL O (1.2.2) multiple sequence alignment

		D31N (human hUTP23)	
<i>A.thaliana</i>	MRVKRQKKNRRTVRFFTVCYGFRQPYKVLCDGTFVHHLVTNEITPADTAVSELLGGPVKL	60	
<i>S.pombe</i>	MRQKRAKNYRKLMHTYQLLFGFREPYQVLVDADFLKDLSSQQKIDIQ-AALARTVQGAIKP	59	
<i>S.cerevisiae</i>	MRQKRAKSYRKQLLVYSHTFKFREPYQVLVDNQLVLECNNSNFNLP-SGLKRTLQADVKV	59	
<i>C.albicans</i>	MRQKRAKAYKKQMSVYVHAFKFREPYQIIVDNELITTCQSASF DIN-KGFTRTIQAENKP	59	
<i>D.melanogaster</i>	MKISRFFKSHKTLVFFASNFDYREPYQVLIDATFCQAALQQKIGID-EQIKKYFQCGVKL	59	
<i>D.rerio</i>	MKIKRQKHAKKTISFYKYNFSFREPFQILIDGTFCAALKNKIQIK-EQLPKYLMGEIQL	59	
<i>M.musculus</i>	MKITRQKHAKKHLGFFRNNGFVREPYQILLDGTFCQAALRGRIQLR-DQLPRYLMGETQL	59	
<i>H.sapiens</i>	MKITRQKHAKKHLGFFRNNGFVREPYQILLDGTFCQAALRGRIQLR-EQLPRYLMGETQL	59	
	*: . * * : : : : * : * : : : * : : : : : :		
		C102A (yeast yUtp23)	
<i>A.thaliana</i>	FTRTCVIAELEKLKGDFAESLEAAQTLNLTATCEHE--EAKTADCECLSEVIG--VQNTHEHF	116	
<i>S.pombe</i>	MITQCCIRQLYSKSDCLKQEIRIAKSFERRRCCHI-DEALSPSECIQSVVNINGRNKHRY	118	
<i>S.cerevisiae</i>	MITQCCIQALYETRND--GAINLAKQFERRRCNHSFKDPKSPAECIESVVNISGANKHRY	117	
<i>C.albicans</i>	MITQCCIQALYDTKNQ--PAIDIAKSFERRRCNHR--EAIDPSQECIESIVNIKGQNKHRY	115	
<i>D.melanogaster</i>	LTTQCVILESESLGAPLTGATSIVKRFHVHKCGHE-GKPVPASECIKSMTK--D--NR	113	
<i>D.rerio</i>	CTTNCALKELSLAKDLYGAKLILQRFQIRKCKHM-KDPVPASECLLSMLA--ETNPHHY	116	
<i>M.musculus</i>	CTTRCVLKELETGKELYGAKLIAQKCQVRNCPHF-KSPVSGSECLLSMVD--EGNPHHY	116	
<i>H.sapiens</i>	CTTRCVLKELETGKLDLYGAKLIAQKCQVRNCPHF-KNAVSGSECLLSMVE--EGNPHHY	116	
	*. * : : : : * * : : * : : : : :		
		C103A (human hUTP23)	
<i>A.thaliana</i>	FLGTQDAEFRKRLQQESIVPLVFLRNILLIDQPSDFQRQSAKDSENKRLTMTDTEKKLL	176	
<i>S.pombe</i>	VVATQDPELRQALRSVPGVPLIYMKRSVILEPASRATLLEKHKNKESVQMGMSKEEKL	178	
<i>S.cerevisiae</i>	VVASQDIDLRRKLRTVPVGPLIHLTRSVVMMEPLSTASAKASKI-----TEEQKLY	168	
<i>C.albicans</i>	IVASQDLQLRKKLRKIPGVPLIYMNRSMVMMEPLSDVSNQYNMN-----YESKKLT	166	
<i>D.melanogaster</i>	VVASQDRLLQESLRLKIPGRCLLYLHKATPVLEAPSKASKKWWQRRAKN-L-----M	163	
<i>D.rerio</i>	FIATQDQQLTTALKKIPGVPLLYIILNTMVLDDKPSERTLKHVEAVQLGEI-----V	167	
<i>M.musculus</i>	FVATQDQNLQSVKVKRTPGIPLMFI IQNTIVLDKPSRPTVAFVKAVEAGQL-----V	167	
<i>H.sapiens</i>	FVATQDQNLQSVKVKKKPGVPLMFI IQNTMVLDDKPSPKTIAFVKAVESGQL-----V	167	
	. : . * * : : : * : : : : *		
<i>A.thaliana</i>	VKRTAKIIASNRKEATIANEEWGMPRVVSTKNGLG-VKDRPQFKRNRAGPNPLSCMKKK	235	
<i>S.pombe</i>	SGKKRSANELAID-DQDTKESTDLA-----GTEDSAPKANKKRGKPGKPNPLSIKKRS	230	
<i>S.cerevisiae</i>	KGLNDPNIE-KLQ---E---SG--D-----GSGKE-SITKKRKLGPAPNPLSVKKKK	211	
<i>C.albicans</i>	GGLNDIEAG-KLE---KQNEGE--D-----GDGDESEVKKKKRKGPKPNPLSVKKKK	213	
<i>D.melanogaster</i>	LQKQVE---KID---YMKEKQGLK-----PAE-TA-VKPKKHKGPKPNPLSCKKSK	207	
<i>D.rerio</i>	NPAQQK---SLQ---SLKEKEGIS-----GDA-EKRG-RKRKRKQSNPNPLSCLKKK	211	
<i>M.musculus</i>	SVHEKQ---SIK---QLKEEQGLV-----RNP-DLRRRRRKKKKVGGPNPLSCLKKK	212	
<i>H.sapiens</i>	SVHEKE---SIK---HLKEEQGLV-----KNT-EQSR-RKKRKKISGPNPLSCLKKK	211	
	. : : ***** *		
<i>A.thaliana</i>	KENPQSKSKADSNSNAQKEKKEGSDTQKRSRKRSSKKGK-SGPORTE-----	281	
<i>S.pombe</i>	SKNHTTDEPTLPVNIIGD-----VGERKKHRRKRK-----	260	
<i>S.cerevisiae</i>	KVNPSDEVKDK----ED----TSKEKKRRRRRKH---KSNTN-----VPVSNGTTAA	253	
<i>C.albicans</i>	TDNATAASTNQE----QK----K---KPNRRKRHG---KSKAEKEDEQEQQVNEATTNE	259	
<i>D.melanogaster</i>	KDKAKQQLKGVET-----AITKAKRKRKIPAHVKAALGKD-----	244	
<i>D.rerio</i>	KKKATPQQPKNP-D-----GEKKRKRSRHRK--HKPAGGEQTEVRS-----	249	
<i>M.musculus</i>	KKAQDTKSP----A-----SEKKRKRKRIRNRSTLKVSSQQGAEG-----	249	
<i>H.sapiens</i>	KKAPDTQSS----A-----SEKKRKRKRIRNRSNPKVLSEKQNAEGE-----	249	
	. . . * : :		
<i>A.thaliana</i>	-----	281	
<i>S.pombe</i>	-----	260	
<i>S.cerevisiae</i>	--Q-----	254	
<i>C.albicans</i>	DAQEAITATE	269	
<i>D.melanogaster</i>	-----	244	
<i>D.rerio</i>	-----	249	
<i>M.musculus</i>	-----	249	
<i>H.sapiens</i>	-----	249	

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Ribosome biogenesis pathways in yeast and humans

Simplified overview of the pre-ribosomal RNA processing pathways in *S. cerevisiae* (A) and *H. sapiens* (B). Cleavages important for 18S rRNA processing are indicated. The positions of radiolabelled probes used for northern blotting (h18SE and hITS1) are marked above the primary transcript. ETS: external transcribed spacer; ITS: internal transcribed spacer.

Figure S2: *In vivo* RNA-protein crosslinking studies to define yUtp23 binding sites

A Outline of the CRAC crosslinking technique (UV crosslinking and analysis of cDNAs).

B Proteins purified from an UV-crosslinked yeast strain expressing C-terminally HTP-tagged yUtp23 were separated by SDS-PAGE and visualised by immunoblotting using the anti-TAP antibody (left panel), which recognises the C-terminus of the HTP construct after TEV cleavage. Crosslinked RNA fragments, which were co-purified with the yUtp23 protein, were radioactively labelled and detected by autoradiography (right panel). Box: membrane cut for isolation of crosslinked RNA.

C Recombinant, GST-tagged yUtp23, yRrp7 or free GST used for electromobility shift assays (EMSA) or protein-protein interaction studies were expressed in *E. coli* and purified using glutathione sepharose. Proteins were separated by SDS-PAGE and stained with Coomassie blue.

D EMSA showing the binding of 5000 nM GST-yUtp23 or GST to trace amounts of an *in vitro* transcribed radiolabelled 18S rRNA ES6 fragment (nt 775-963, RNA A). Binding was performed in the absence (lanes 1 and 4) or with an excess (3 pmoles) of non-radiolabelled competitor RNA A (18S nt 775-963, lane 2) or RNA B (18S nt 1022-1146, lane 3). RNP complexes were separated from unbound RNA on 4% native polyacrylamide gels and visualised using a PhosphorImager. RNP levels from three independent experiments were determined by quantification using ImageQuant and normalised to the reaction containing no competitor RNA (lane 1).

Figure S3: Primer extension analysis of yUtp23/hUTP23 crosslinking sites in the yeast and human 18S ES6 region

A RNA-containing complexes were isolated from yeast strains expressing C-terminally HTP-tagged yUtp23, or the exosome component yRrp44 as a control. Yeast cells were either crosslinked (100 sec at 254 nm) (+UV) or non-treated (-UV), and purifications were performed as described for CRAC, with the exception that purified complexes were not treated with RNase. Briefly, complexes were first purified on IgG sepharose, eluted by TEV cleavage and re-purified on Nickel agarose under denaturing conditions (1,2). Co-purified RNA was extracted and analysed by primer extension using a primer hybridising to nt 943-

963 of the mature yeast 18S rRNA (left panel). *In vitro* transcribed RNA mimicking the yeast 23S rRNA was used to generate a sequencing ladder. Reverse transcriptase stops that are only present on RNA isolated from crosslinked yUtp23 complexes are marked by circles, a stop also present in the crosslinked control sample (yRrp44) is marked by an asterisk. Right panel: Predicted structure of the yeast 18S ES6 region after snR30 release, highlighting the positions of the primer (shaded in grey) and the reverse transcriptase stops shown in the left panel (circles and asterisk).

B RNA-containing complexes were isolated from crosslinked (+UV) or non-treated (-UV) HEK293 cells expressing C-terminally His₈-PP-2xHA-tagged wild type hUTP23. UV crosslinking was performed in the Stratalinker (3 x 800 mJ/cm²) and purifications were performed from sonicated whole-cell extracts as described for CRAC (1,2), with the exception that purified complexes were not treated with RNase. Briefly, complexes were first purified on anti-HA agarose, eluted by PreScission protease cleavage and then re-purified on Nickel agarose under denaturing conditions. Co-purified RNA was extracted and analysed by primer extension using a primer hybridising to nt 1001-1026 of the mature human 18S rRNA (left panel). Total RNA extracted from HEK293 cells was used to generate a sequencing ladder. Reverse transcriptase stops visible in all lanes except the non-crosslinked hUTP23 sample are likely due to secondary structure of the 18S rRNA and marked by asterisks. Circle: Reverse transcriptase stop that is only present in RNA isolated from crosslinked hUTP23 complexes. Right panel: Predicted structure of the human 18S ES6 region after U17 release, highlighting the positions of the primer (shaded in grey) and the reverse transcriptase stops shown in the left panel (circle and asterisks).

Figure S4: yUtp23 crosslinking sites on small nucleolar RNAs (snoRNAs)

A Yeast and human snoRNAs with non-modifying roles important for A0, A1/1 and A2/2a cleavages

B Top 20 snoRNA hits. Sequence data from two individual yUtp23 datasets are shown as percentages of all reads mapped to snoRNAs.

C yUtp23 crosslinking sites on three snoRNAs involved in SSU production are plotted as total reads (hits) or microdeletions (dels) per million mapped sequences. Top: SNR17A/B (U3); middle: SNR10 and bottom: SNR128 (U14). Data from the two individual datasets shown in panel **B** were combined.

D Predicted secondary structure of the human U17 box H/ACA snoRNA. The H and ACA boxes are shown in black.

Figure S5: Recombinant proteins used for protein-protein interaction and RNA binding assays

Recombinant, GST-tagged wild type or mutant Utp23 proteins from human (hUTP23, **A**) or yeast (yUtp23, **B**) or the GST tag alone (**B**) were expressed in *E. coli* and purified using glutathione sepharose. Proteins were separated by SDS-PAGE and stained with Coomassie blue. D31N: hUTP23 PIN mutant. C103A: hUTP23 Zinc finger mutant. C102A: yUtp23 Zinc finger mutant.

Figure S6: Utp23 protein alignment

Selected plant, fungal and metazoan protein sequences retrieved by a NCBI protein BLAST search using the *S. cerevisiae* yUtp23 (YOR004W) protein sequence were aligned with clustalO (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). The alignment was modified to highlight acidic residues (red) within the PIN domain predicted catalytic centre (shaded in yellow) and conserved residues in the CCHC Zinc finger motif (blue), respectively.

SUPPLEMENTARY TABLES

Table S1: Yeast strains used in this study

BY4741	<i>MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0</i>	(Euroscarf)
yUtp23-HTP	<i>MATa; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; UTP23-HTP-URA3</i>	(this study)

Table S2: Plasmids used in this study

Insert	Vector	Purpose	
<i>Yeast</i>			
GST-yUtp23-WT	pGEX-6P1	To purify recombinant protein	(this study)
GST-yUtp23-C102A	pGEX-6P1	To purify recombinant protein	(this study)
GST-yUtp24	pGEX-6P1	To purify recombinant protein	(3)
GST-yRrp7	pGEX-6P1	To purify recombinant protein	(this study)
ProtA-TEV-yUtp23-His ₆	Nzztev80 (4)	To purify recombinant protein	(this study)
His ₆ -yUtp23	pET100	To produce <i>in vitro</i> translates	(this study)
His ₆ -yRok1	pET100	To produce <i>in vitro</i> translates	(this study)
His ₆ -yDhr1	pET100	To produce <i>in vitro</i> translates	(this study)
His ₆ -yKri1	pET100	To produce <i>in vitro</i> translates	(this study)
His ₆ -yNhp2	pET100	To produce <i>in vitro</i> translates	(this study)
His ₆ -yRrp9	pET100	To produce <i>in vitro</i> translates	(this study)
pT7snR30 (cut BamHI)	pSP6T7	For <i>in vitro</i> transcription (T7 Pol)	(5)
<i>Human</i>			
hUTP23-WT-His ₈ -PP-2HA	pcDNA5	To stably transfect HEK293 cells	(this study)
hUTP23-D31N-His ₈ -PP-2HA	pcDNA5	To stably transfect HEK293 cells	(this study)
hUTP23-C103A-His ₈ -PP-2HA	pcDNA5	To stably transfect HEK293 cells	(this study)
GST-hUTP23-WT	pGEX-6P1	To purify recombinant protein	(this study)
GST-hUTP23-D31N	pGEX-6P1	To purify recombinant protein	(this study)
GST-hUTP23-C103A	pGEX-6P1	To purify recombinant protein	(this study)
His ₆ -hUTP24	pET100	To purify recombinant protein	(this study)
His ₆ -hRRP7	pET100	To purify recombinant protein	(this study)
His ₆ -hNHP2	pET100	To purify recombinant protein	(this study)
His ₆ -hROK1/DDX52	pET100	To produce <i>in vitro</i> translates	(this study)
pBS+HU17	pBSIISK+	To amplify U17 for transcription with T7 Pol (see Table S3)	(6)

Table S3: Oligonucleotides (5'-3') used in this study

Strain construction	
yUtp23-HTP	
F	GCAACACTAACGTCCCCGTTTCAAATGGGACCACAGCCGCGCAG GAGCACCATCACCATCACC
R	CGCAATTATATTCAAAAAGATTGGCCAGTCAATTAAGGAAAA AATTACGACTCACTATAGGG
CRAC	
3'-Linker	5'- rAppTGGAATTCTCGGGTGCCAAGG/ddC/ -3' (miRCat-33™)
5'-Linker (L5Aa)	5'-invddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrU rNrNrNrTrArArGrC-OH-3'
5'-Linker (L5Ac)	5'-invddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrU rNrNrNrGrCrGrCrArGrC-OH-3'
5'-Linker (L5f)	5'-invddT-ACACrGrArCrGrCrUrCrUrUrCrCrGrArUrCrU rGrCrGrArGrC-OH-3'
RT	CCTTGGCACCCGAGAATT (miRCat-33™)
PCR - F	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGAC GCTCTTCCGATCT
PCR - R	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGGCC TTGGCACCCGAGAATTCC
Cloning (Vector)	
<i>Yeast</i>	
yUtp23 (pET100 and pGEX-6P1)	
F (BamHI)	CACCGGATCCATGCGTCAAAAGAGGGCTAAGTCATATAGG
R (Sall)	GTCGACTTACTGCGCGGCTGTGGTC
yUtp23 (Nzztev80) (4)	
F (SphI)	TTCAGGGCATGCTGGAAGTTCTGTTCCAGGGACCCCGTCAAAAG AGGGCTAAGTCATATAG
R (XmaI)	CGTCGACCCGGGGCTGCGCGGCTGTGGTC
yRrp7 (pGEX-6P1)	
F (BamHI)	GCGCGGATCCATGGGTATTGAAGACATTAGCGC
R (Sall)	CCCCGTCGACTTAAGTGTATGGATTGAATTTTCTCTTAGC
yRok1 (pET100)	
F (BglII)	CACCAGATCTATGGATATTTTATAGATATTAAGTAGAGGAGC
R (XhoI)	CTCGAGTTATTTTCGAGAAATGTTTTTTTGAAG
yDhr1 (pET100)	
F (BglII)	CACCAGATCTATGGGTACTTACAGAAAAAGGTTTAATG
R (XhoI)	CTCGAGTTATTTTTTTCTTTCTTTCACCTGTG
yKri1 (pET100)	
F (BglII)	CACCAGATCTATGCCAAGAAAAAAGTCTGCC
R (XhoI)	CTCGAGTTACTTTTTCTGGTGGCCTTTATG
yNhp2 (pET100)	
F (BamHI)	CACCGGATCCATGGGTAAAGACAACAAGGAACATAAG
R (Sall)	GTCGACTCATAAAGCTTGAAGTTCTTTGACAAC

yRrp9	(pET100)
F (BamHI)	CACCGGATCCATGTCAGATGTTACCCAACAGAAAAAGAG
R (Sall)	GTCGACCTAAAAGCCTGTTTGGTCAATGACAG
<i>Human</i>	
hUTP23	(pGEX-6P1)
F (BamHI)	CACCGGATCCATGAAGATCACAAGGCAGAAACATGCC
R (XhoI)	CGCGCTCGAGTCATTCTCCTTCTGCATTCTGCTTCTC
hUTP24	(pET100)
F (BglII)	CACCAGATCTATGGGGAAGCAAAAGAAAACAAGGAAGTATGCGA CCATGAAGCGAATGCTTAG
R (XhoI)	CGCGCTCGAGTTAGAATCGAGGGGCTCCATAATCATCTGG
hRRP7	(pET100)
F (BamHI)	CACCGGATCCATGGTGGCGCGCAGGAGGAAGTGCGC
R (XhoI)	CGCGCTCGAGTCAGTACGGTCGGAATTTGCGC
hNHP2	(pET100)
F (BamHI)	CACCGGATCCATGACCAAATAAAGGCAGATCC
R (XhoI)	CTCGAGTTAAAGGGGTAGGGGCAGG
hROK1/DDX52	(pET100)
F (BamHI)	CACCGGATCCATGGACGTCCACGATCTCTTTC
R (XhoI)	CTCGAGTTAACTTTTGTCTTCAAGAGCTACTTTC
Mutagenesis	(Mutation)
<i>Yeast</i>	
yUtp23	C102A (Zinc finger mutant)
F	CAAAGTCACCTGCGGAGGCCATCGAAAGCGTCGTAA
R	TTAACGACGCTTTCGATGGCCTCCGCAGGTGACTTTG
snR30 ΔIH	To delete the IH of snR30 (nt 398-531)
F	GTTTAACTTAGATTAAGCCGCAGTATATTCCTAAACACTATGAAAT GAC
R	GGAATATACTGCGGCTTAATCTAAGTTAACTCGTCAACGGGG
<i>Human</i>	
hUTP23	D31N (PIN mutant)
F	CCGTACCAGATCCTGCTGAACGGCACCTTCTGTCAG
R	CTGACAGAAGGTGCCGTTACGACAGGATCTGGTACGG
hUTP23	C103A (Zinc finger mutant)
F	GCAGTGAGTGGATCAGAAGCTCTGCTTTCCATGGTTGAAGAG
R	CTCTTCAACCATGGAAAGCAGAGCTTCTGATCCACTCACTGC
Northern blotting	
hITS1 (6121)	AGGGGTCTTTAAACCTCCGCGCCGGAACGCGCTAGGTAC
h18SE	CCTCGCCCTCCGGGCTCCGTTAATGATC
U17	TTCCTGCATGGTTTGTCTCC
7SK	GTGTCTGGAGTCTTGGAAGC
Primer extension	
y18S A963R	TGAAAACGTCCTTGGCAAATG
h18S A1026R	TGAAAACATTCTTGGCAAATGCTTTC

For transcription	
	(yeast 18S rRNA: nt 775-963)
F	<u>CGGAATTCTAATACGACTCACTATAGG</u> CGTATTGCTCGAATATATT AGC (<u>T7 promoter</u>)
R	TGAAAACGTCCTTGGCAAATG
	(yeast 18S rRNA: nt 1022-1146)
F	<u>CGGAATTCTAATACGACTCACTATAGG</u> CATAAACTATGCCGACTA GGGATC (<u>T7 promoter</u>)
R	CAATTCCTTTAAGTTTCAGCCTTG
	(yeast 23S rRNA: nt 1-2786)
F	<u>CGGAATTCTAATACGACTCACTATAGG</u> ATGCGAAAGCAGTTGAAG AC (<u>T7 promoter</u>)
R	GTTTGTTACCTCTGGGCCCC
	(snR30 fragments)
F (nt 1)	<u>CGGAATTCTAATACGACTCACTATAGG</u> AACCATAGTCTCGTGCTA GTTCGG (<u>T7 promoter</u>)
F (nt 398)	<u>CGGAATTCTAATACGACTCACTATAGG</u> GTAGGACGCATGATCTTG AGCTC (<u>T7 promoter</u>)
F (nt 531)	<u>CGGAATTCTAATACGACTCACTATAGG</u> CCGCAGTATATTCCTAAAC ACTATG (<u>T7 promoter</u>)
R (nt 398)	CTTAATCTAAGTTAAACTCGTCAACGG
R (nt 531)	GTAGGACGAACAACAAAGATGACC
R (nt 606)	AGATGTCTGCAGTATGGTTTTACCC
	(U17 snoRNA)
F	<u>GCGTAATACGACTCACTATAGGG</u> CGATCCAACGTGGATACACCC GG (<u>T7 promoter</u>)
R	GGCTGTTTCCTGCATCGGTTTGTC

Table S4: siRNAs used in this study

Gene	Target Sequence (5'-3')	Source
Luciferase GL2	CGUACGCGGAUACUUCGA	Eurofins MWG (7)
hUTP23	GAAAGUAUCAACAUCUCATT	Eurofins MWG

Table S5: Antibodies used in this study

Target	Source	Cat. No.	Dilution	Species
Western blotting				
α -hKaryopherin	Santa Cruz	sc-11367	1:1000	rabbit
α -hUTP23	Aviva	ARP60526_P050	1:1000	rabbit
α -HA	Babco	MMS-101P	1:5000	mouse
α -TAP*	Thermo Scientific	CAB1001	1:10,000	rabbit
α -mouse-HRP	Santa Cruz	sc-2316	1:10,000	donkey
α -rabbit-HRP	Santa Cruz	sc-2313	1:10,000	donkey
α -rabbit-IRDye 800CW	Li-Cor	926-32211	1:10,000	goat
Immunofluorescence				
α -HA	Babco	MMS-101P	1:500	mouse
α -hFibrillarin	Santa Cruz	sc-25397	1:200	rabbit
α -rabbit-IgG (Alexa Fluor 555 conjugate)	Invitrogen	A-31572	1:500	donkey
α -mouse-IgG (Alexa Fluor 555 conjugate)	Invitrogen	A-31570	1:500	donkey

* The α -TAP antibody recognises the C-terminus of the HTP construct before and after TEV cleavage.

SUPPLEMENTARY REFERENCES

1. Granneman, S., Kudla, G., Petfalski, E. and Tollervey, D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A*, **106**, 9613-9618.
2. Granneman, S., Petfalski, E. and Tollervey, D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J*, **30**, 4006-4019.
3. Wells, G.R., Weichmann, F., Colvin, D., Sloan, K.E., Kudla, G., Tollervey, D., Watkins, N.J. and Schneider, C. (2016) The PIN domain endonuclease Utp24 cleaves pre-ribosomal RNA at two coupled sites in yeast and humans. *Nucleic Acids Res*, **44**, 5399-5409.
4. Mingot, J.M., Bohnsack, M.T., Jakle, U. and Gorlich, D. (2004) Exportin 7 defines a novel general nuclear export pathway. *EMBO J*, **23**, 3227-3236.
5. Watkins, N.J., Gottschalk, A., Neubauer, G., Kastner, B., Fabrizio, P., Mann, M. and Luhrmann, R. (1998) Cbf5p, a potential pseudouridine synthase, and Nhp2p, a putative RNA-binding protein, are present together with Gar1p in all H BOX/ACA-motif snoRNPs and constitute a common bipartite structure. *RNA*, **4**, 1549-1568.
6. Watkins, N.J., Leverette, R.D., Xia, L., Andrews, M.T. and Maxwell, E.S. (1996) Elements essential for processing intronic U14 snoRNA are located at the termini of the mature snoRNA sequence and include conserved nucleotide boxes C and D. *RNA*, **2**, 118-133.
7. Elbashir, S.M., Harborth, J., Weber, K. and Tuschl, T. (2002) Analysis of gene function in somatic mammalian cells using small interfering RNAs. *Methods*, **26**, 199-213.